VOL. 16 No. 5-6, pp. 141-180 NOVEMBER-DECEMBER 1954

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

FERROXDURE II AND III, ANISOTROPIC PERMANENT MAGNET MATERIALS

by A. L. STUIJTS, G. W. RATHENAU *) and G. H. WEBER. 621.318.124:538.246.2

*The $(BH)_{\max}$ product, i.e. the maximum value of the product of the induction B and the magnetizing force H for points on the demagnetization curve, is a measure of the quality of any permanent magnet material. In Ferroxdure **), a ceramic permanent magnet material already discussed in this Review, values of $(BH)_{\max}$ up to about 7200 J/m³ [0.9×10^6 gauss oersted] are attained. This article describes the development of an anisotropic type of Ferroxdure, which has a $(BH)_{\max}$ product comparable to that of modern magnet steels, viz. up to about 28 000 J/m³ [3.5×10^6 gauss oersted].*

Characteristic properties of Ferroxdure

Ferroxdure **) is the name given to a new group of ceramic permanent magnet materials, some properties of which have been fully described in an earlier issue of this Review [1]). The magnetically essential phase of Ferroxdure I (which is now in quantity production) is the compound $BaFe_{12}O_{19}$, whose crystal structure is hexagonal, with close-packed oxygen ions. One of the oxygen sites in this structure is occupied by a Ba^{2+} ion; the Fe^{3+} ions occupy five non-equivalent lattice sites in the interstices between the oxygen ions.

The saturation magnetization of the compound $BaFe_{12}O_{19}$ is low compared with that of iron; extrapolation to absolute zero gives $J_0 = 0.66$ Wb/m² [$4\pi J_0 = 6600$ gauss] [2]) as against 2.18 Wb/m² [21 800 gauss] for iron. Again, at room temperature the value J_s for this compound is 0.475 Wb/m² [$4\pi J_s = 4750$ gauss], whereas the corresponding value for iron is 2.15 Wb/m² [21 500 gauss]. This relatively low saturation value and also the variation of the paramagnetic susceptibility as a function of temperature, suggest that the source of magnetization in the compound is what may be described as non-compensated antiferromagnetism, or ferrimagnetism [3]). Although the magnetic moment of each Fe^{3+} ion in this ionic compound is more than twice the moment per atom of metallic iron, only one third of the total moment of these ions is really effective, since the residual elementary moments are oriented antiparallel to each other and therefore cancel. Moreover, the magnetization is all the lower in relation to that of iron owing to the fact that the structure of the compound is greatly "diluted" with oxygen ions.

The strong coercive force exhibited by Ferroxdure is attributable to a pronounced magnetic anisotropy, in this case mainly crystal anisotropy, as will now be explained. In every Ferroxdure crystal there exists one direction of easy magnetization (preferred direction), parallel to the hexagonal axis, and only a strong magnetic field can turn the magnetization away from this preferred direction. It is to this distinct crystal anisotropy, combined with a relatively low saturation magnetization, that Ferroxdure owes its magnetic hardness. As a result of the low saturation magnetization, the couple

$$M = J_s H \sin \varepsilon, \quad . \; . \; . \; . \; . \quad (1)$$

which a magnetic field at an angle ε to the preferred direction can exert to rotate the magnetization away from this direction is relatively weak, whereas the couple

---

*) Now professor at the University of Amsterdam; formerly with the Philips Laboratories at Eindhoven.

**) Known in U.S.A., U.K., and some other countries under the trade name "Magnadur".

[1]) J. J. Went, G. W. Rathenau, E. W. Gorter and G. W. van Oosterhout, Ferroxdure, a class of new permanent magnet materials, Philips tech. Rev. **13**, 194-208, 1951/52.

[2]) Rationalized Giorgi-units are employed throughout this article (see Philips tech. Rev. **10**, 55-60, 1948/49 and **13**, 194, 1951/52. For convenience, the formulae and values as expressed in unrationalized cgs-units (Gauss units) are also quoted, in square brackets.

[3]) See article referred to in note [1]).

$$M = K \sin 2\Theta \quad . \; . \; . \; . \; . \; . \quad (2)$$

tending to turn the magnetization back to the preferred direction is strong (K is the anisotropy constant and Θ the angle which the direction of of magnetization in a field H makes with the preferred direction).

Given this process of rotation alone, the coercive force, as computed for a material in which the preferred directions of the constituent crystals are distributed at random, is:

$$_JH_c = 0.96\, K/J_s \,.$$

For Ferroxdure at room temperature, $K = 3.1 \times 10^5$ J/m^3 [3.1×10^6 erg/cm^3] and $J_s = 0.475$ Wb/m^2 [$4\pi J_s = 4750$ gauss]: hence a value of $\mu_0 {_JH_c} = 0.79$ Wb/m^2 [$_JH_c$ = 7900 oersted] is calculated. In practice, however, a value of about 0.3 Wb/m^2 [3000 oersted] is obtained in the case of Ferroxdure I, which suggests that Bloch wall movements, which occur far more readily, are also involved. Now, no stable Bloch walls can be formed in particles of diameter less than a certain critical value. In the case here considered, the critical diameter is about 1μ, which, again by reason of the distinct crystal anisotropy and relatively low saturation magnetization of the material, is considerably larger than the critical diameter for iron and cobalt crystals. Accordingly, an effective sintering process, producing a compact material whose constituent particles are smaller than the critical diameter, is the first essential. In practice, however, the sintered product will invariably contain a number of larger particles, which tend to diminish the coercive force by fostering the formation and movement of Bloch walls.

Permanent magnets are often evaluated in terms of the $(BH)_{\max}$ product, that is, the maximum value of the product of the induction B and the field H for points on the demagnetization curve. The volume of material required to maintain a given magnetizing force in a particular air-gap, broadly speaking, is inversely proportional to the $(BH)_{\max}$ product. Hence there is every reason to increase the value of this product as far as possible. The formula for an "ideally permanent" magnet ($_JH_c = \infty$, $\mu_0 {_BH_c} = B_r$) is (see *fig. 1*):

$$\left.\begin{array}{l} (BH)_{\max} = \dfrac{1}{\mu_0}\left(\dfrac{B_r}{2}\right)^2 \\ \text{or,} \quad \left[(BH)_{\max} = \left(\dfrac{B_r}{2}\right)^2\right] \end{array}\right\} \quad . \; . \; . \; . \; . \quad (3)$$

In the case of Ferroxdure I, $\mu_0 {_JH_c} \gg B_r/2$ [$_JH_c \gg B_r/2$], which allows the above formula to hold reasonably well, as will be seen from the fact that the theoretical maximum is:

$$(BH)_{\max} \approx \frac{1}{\mu_0}\left(\frac{0.21}{2}\right)^2 \approx 8800 \text{ J/m}^3$$

$$\text{or,} \quad \left[(BH)_{\max} \approx \left(\frac{2100}{2}\right)^2 \approx 1.1 \times 10^6 \text{ gauss Oe}\right],$$

as compared with 7200 J/m^3 [0.9×10^6 gauss oersted], which is the value established by experiment.

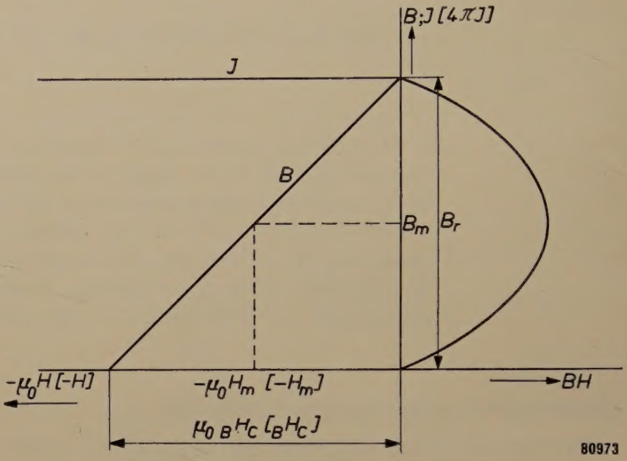


Fig. 1. Demagnetization curve for a material of which $_JH_c = \infty$. In the case here considered,

$$(BH)_{\max} = \frac{1}{\mu_0}\left(\frac{B_r}{2}\right)^2, \quad \text{or} \quad \left[(BH)_{\max} = \left(\frac{B_r}{2}\right)^2\right].$$

Whereas a low remanence limits the value of $(BH)_{\max}$ in Ferroxdure I, the fact that $\mu_0 {_JH_c} \ll B_r/2$ [$_JH_c \ll B_r/2$] is mainly responsible for the limitation of this product in *metallic* magnets.

Ferroxdure I is very suitable for use in cases where a high coercive force is required, that is, where the material is exposed to very strong demagnetizing fields, as in focusing magnets for television tubes [4]), magnets for bicycle dynamos, yoke magnets and magnetic clutches. By virtue of its high resistivity, this material can be employed for producing a biasing field in high-frequency transformer cores of Ferroxcube [5]). Other advantages of Ferroxdure are its simple composition and low density, and also the fact that it is made of relatively inexpensive and readily obtainable materials. Moreover, the magnetization loss caused by demagnetizing fields is limited by the low permeability of this material.

The use of Ferroxdure I to sustain a strong magnetizing force in a narrow air-gap, as in a loudspeaker circuit, is confined mainly to those cases which favour the use of a very flat magnet. However, a complete change in the situation, particularly in this respect, has taken place as a result of a new development [6]), which has led to the production of Ferroxdure materials showing an appreciably higher induction on the demagnetization curve, and having $(BH)_{\max}$ values comparable with those of modern anisotropic magnet steels.

---

[4]) J. A. Verhoef, Philips tech. Rev. **15**, 214-220, 1953/54 (No. 7).

[5]) W. Six, Philips tech. Rev. **13**, 301-311, 1951/52.

Anisotropic Ferroxdure

In the isotropic material, the preferred directions of magnetization of the constituent particles are oriented at random. It is found that the remanence B_r of material so constituted is equal to half the saturation magnetization, i.e. about 0.21 Wb/m² [2100 gauss]. Given a material in which the preferred directions of the individual particles are aligned parallel to one another, however, the remanence in this direction will be the same as in a single crystal, viz.:

$$B_r = J_s, \qquad [B_r = 4\pi J_s].$$

Accordingly, we have as a maximum:

$$(BH)_{\max} \approx \frac{1}{\mu_0}\left(\frac{0.42}{2}\right)^2 \approx 35\,000 \text{ J/m}^3$$

$$\left[(BH)_{\max} \approx \left(\frac{4200}{2}\right)^2 \approx 4.4\times10^6 \text{ gauss oersted}\right].$$

To procure this optimum value, it is necessary to ensure that in all circumstances: $\mu_0 {}_JH_c > B_r/2$ $[{}_JH_c > B_r/2]$. The fulfilment of this condition becomes more and more difficult as the remanence B_r increases. Moreover, the coercivity is invariably weaker in anisotropic than in isotropic material, owing to the fact that the changes in the direction of magnetization are caused by the formation of Bloch walls. The field H^* required to form a Bloch wall in a particle is a minimum when the preferred direction of magnetization of the particular particle coincides with the direction of the applied field H; also, the movement of Bloch walls across grain boundaries takes place far more readily in anisotropic materials, which in this respect resemble single crystals.

Accordingly, the problem is to procure an anisotropic orientation of the grains in Ferroxdure, at the same time avoiding any appreciable weakening of the coercive force.

Such anisotropy can be induced in several ways. For example, one approach to the problem is to take advantage of the non-spherical shape of the Ferroxdure single crystals. These crystals exhibit very strong growth along the basal plane, but relatively slight growth in the hexagonal direction coinciding with the preferred direction of magnetization. By packing the powder in a steel tube, welding the ends of this tube and passing it through rolls at a high temperature, a preferred direction of magnetization at right-angles to the plane of rolling is obtained. Similarly, it is possible to produce a small amount of anisotropy merely by compressing the Ferroxdure powder under high pressure.

However, a far superior method is to employ the magnetocrystalline anisotropy of Ferroxdure to the best advantage. When an external magnetic field H is applied, (*fig. 2*), the preferred axes of magnetization of any particles which are free to rotate will move into alignment with this field. This produces in the material a texture corresponding approximately to the magnetic state of a single crystal, that is, one preferred direction, parallel to the direction of the field.

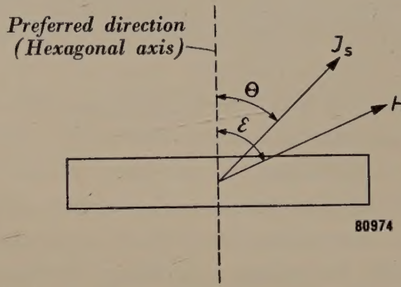


Fig. 2. The state of equilibrium of the magnetization J_s in a uniaxial magnetic particle when an external field H is applied at an angle ε to the preferred direction.

The aligning couple exerted by the magnetization J_s on the crystal is:

$$M = K \sin 2\Theta. \quad \dots\dots \quad (2)$$

The absolute value of this couple reaches a maximum equal to K when $\Theta = 45°$ or 135°, and is zero when $\Theta = 0°$, 90° or 180°.

It will be seen, then, that the highest gain in $(BH)_{\max}$ is attained when those particles whose hexagonal axes make an angle ε of about 90° with the field direction are also brought into alignment. If at $\varepsilon = 90°$ the magnetizing force exceeds $2\,K/J_s$, the magnetization will rotate in the direction of H $(\Theta = 90°)$, the orientation of the crystal axis remaining unchanged; the mechanical couple is then zero. The rotation of the magnetization is instantaneous; hence the application of the field does not usually produce a simultaneous rotation of the particle. Accordingly, the latter will also be able to remain out of alignment when the magnetizing force is increased, However, if H be small enough, e.g. $H = \sqrt{2}K/J_s$, a sudden application of the field will produce the state corresponding to $\Theta = 45°$, i.e. that with the maximum couple: hence the particle will also be brought into alignment with the field. This suggests that the value of an external magnetic field suddenly applied should be approximately equal to $\mu_0 H = \mu_0 \sqrt{2}\, K/J_s \approx 1.1\ \mathrm{Wb/m^2}$ [$H \approx 11000$ oersted].

---

[6]) G. W. Rathenau, J. Smit and A. L. Stuijts. Z. Phys. **133**, 250, 1952. A. L. Stuijts, G. W. Rathenau and E. W. Gorter, J. appl. Phys. **23**, 1282, 1952.

Method of preparation

To apply in practice the process considered here, it is necessary to fulfil a number of conditions.

Firstly, the particles subjected to the field treatment must be magnetic: the basic material is isotropic Ferroxdure. Secondly, the basic material must consist of separate particles, each having only one crystal orientation; this condition can be fulfilled by grinding the sintered material until it contains no aggregates comprising more than one crystal.

A magnetic particle introduced into a magnetic field will turn to the particular direction most favourable from a magnetic point of view, if able to overcome the local forces opposing this orientation. To procure optimum efficiency, then, it is necessary to minimize the tendency of one particle to impede the orientation of another. Since the flatness of Ferroxdure crystals is unfavourable in this respect, too close a packing of the powder subjected to the magnetic field treatment has to be avoided.

When once the particles in the magnetic field are properly aligned, they must be stabilized without any disturbance of the texture thus obtained. By mixing the powder with a little melted paraffin wax and introducing this sludge into the magnetic field, stabilization is obtained when the wax solidifies. However, this method is of little practical value owing to the fact that the induction B is proportional to the density of the magnetic material, which would be low. In the case of ceramic materials, a product combining stability with adequate density is normally obtained by press-forming the powder at high pressure and then sintering it. Experience has shown that pressing is an excellent method of stabilizing the particles, provided that it takes place in a magnetic field; although the forces exerted during this process far exceed the aligning force of the magnetic field, they do not seriously disturb the texture of the material. As regards sintering, the condition that undue particle growth must be avoided in order to maintain a high coercive force is even more important than in the case of Ferroxdure I.

Finally, the high temperature employed in the sintering process, which is well above the Curie point, must not spoil the magnetic texture of the material. This is a point to which we shall shortly return.

The powder is press-formed as an aqueous sludge. A diagram of the apparatus employed is seen in *fig. 3*. The die is made of non-magnetic material, and iron pole pieces are provided to intensify the field within the powder. The surplus liquid is ejected along the plunger. Experience has shown that the relative directions of the compressing force and the magnetic field are immaterial.

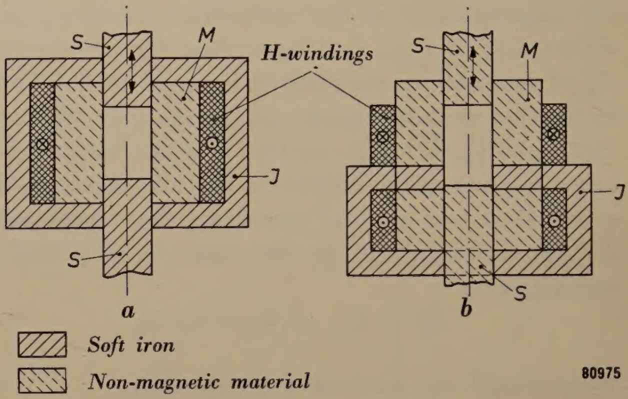


Fig. 3. Diagram of the dies M used to compress the powder in a magnetic field. S = plunger. J = iron yoke.
a) Direction of compression // to direction of field.
b) Direction of compression ⊥ to direction of field.

After pressing, the powder still contains a large amount of water, and must be dried thoroughly before sintering. The technique of press-forming such an aqueous sludge differs considerably from that employed for dry powder and, in fact, the adaption of the process to large-scale production involved the solving of many technological problems.

The effect of sintering

It will be seen, then, that the texture of the compressed powder is not ideal; the particles obstruct one another during alignment, compression affects the texture, and the aligning couple decreases with angle ε (see fig. 2).

It is also necessary that the texture be preserved during sintering. However, the notable feature of this process is that it not only preserves, but actually enhances, the texture. Consider *table I*, which shows the remanence values of a particular material as measured parallel and at right-angles to the preferred direction, after sintering at different temperatures. In the case of a fully anisotropic material

Table I. Remanences of anisotropic Ferroxdure as measured parallel and at right-angles to the direction of the field in which the crystals are oriented, after sintering at different temperatures.

Sintering temp.	Density		Remanence $B_{r//}$ // to original field-direction		Remanence $B_{r\perp}$ ⊥ to original field-direction		$\frac{B_{r\perp}}{B_{r//}}$	Coercivity	
								$\mu_0 {}_JH_c$	${}_JH_c$
°C	$\frac{kg}{m^3}$	$\frac{g}{cm^3}$	$\frac{Wb}{m^2}$	gauss	$\frac{Wb}{m^2}$	gauss		$\frac{Wb}{m^2}$	oersted
1250	4060	4.06	0.270	2700	0.121	1210	0.45	0.321	3210
1275	4360	4.36	0.292	2920	0.125	1250	0.43	0.285	2850
1300	4660	4.66	0.326	3260	0.117	1170	0.36	0.247	2470
1320	4840	4.84	0.362	3620	0.072	720	0.20	0.117	1170
1340	4880	4.88	0.385	3850	0.025	250	0.07	<0.020	<200

$B_{r\perp}/B_{r//} = 0$, whereas in that of the isotropic material $B_{r\perp}/B_{r//} = 1$.

From the magnetic point of view, the texture changes from an imperfect one to that of a single crystal after the Ferroxdure has been sintered at a high temperature. *Fig. 4* shows part of the hysteresis loops of such a material, again measured parallel and at right-angles to the preferred direction. *Figures 5a* and *5b* are photo-micrographs showing the surfaces of a cube of the same material. Note the distinct texture produced by sintering at a high temperature. In the upper photograph the direction of the field applied during compression is at right-angles to the plane of the paper; in the lower photograph it is parallel.

This improvement of textural alignment during sintering could be explained by supposing the small particles in the material to be eliminated in favour of the larger ones by the action of interfacial tensions at the crystal boundaries during the sintering process; this explanation, if correct, implies that the smaller particles are not properly aligned.

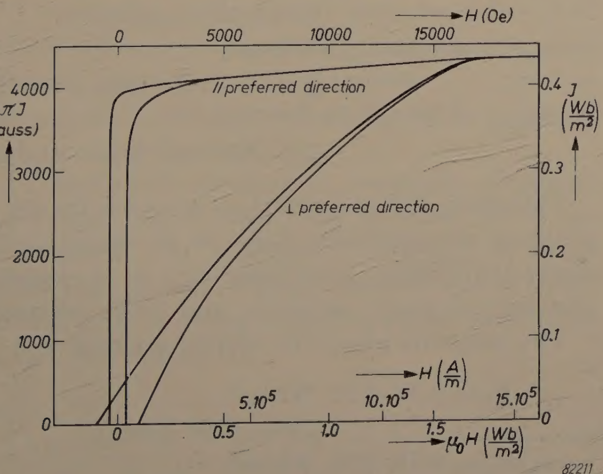


Fig. 4. Hysteresis loops (upper half only) of an isotropic Ferroxdure sample sintered at a high temperature, as measured // and ⊥ to the preferred direction. The magnetic texture closely resembles that of a single crystal.

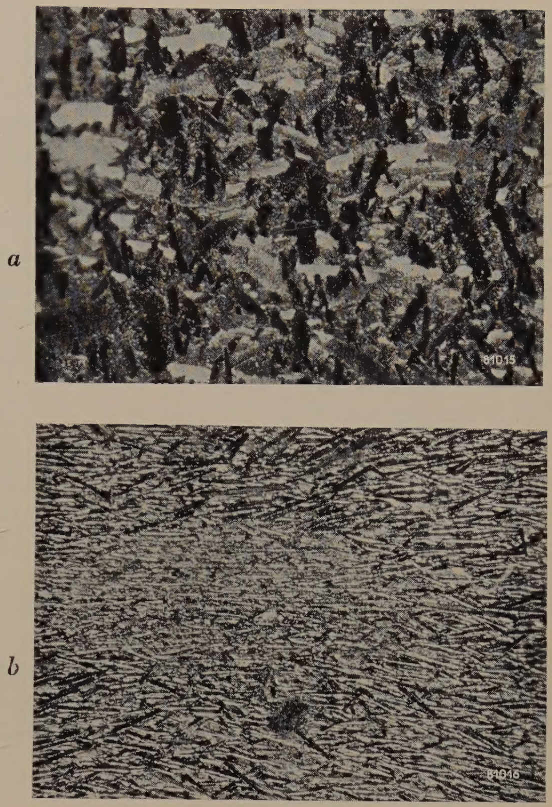


Fig. 5. Photo-micrograph of Ferroxdure with crystals aligned. In (*a*) the hexagonal basal planes are in the plane of the photograph, and in (*b*) they are at right angles to this plane. Magnification 10×.

However, experience has shown that the most significant improvement of the texture takes place when the density of the material is already about 90% of that of an ideal single crystal; it will be seen from the table that the coercive force then decreases very appreciably.

Accordingly, the following explanation is in our opinion more feasible, although, no doubt, the above-mentioned mechanism is also involved.

In the case of metals, the beneficial effect of grain growth on a preferred orientation has been closely studied. This effect can be observed direct with the aid of an emission electron-microscope [7]). It is seen that appreciable grain growth takes place only when the difference in orientation between adjacent crystals is such that the boundary energy between these crystals is high. Given a matrix of identically, or almost identically, oriented crystals surrounding one crystal differently aligned, this odd crystal will usually be eliminated, whereas the boundaries between the similarly aligned crystals will be preserved. Again, it is found that the odd crystal can grow only if it is large in relation to the crystals of the matrix.

---

[7]) G. W. Rathenau and G. Baas, Physica **17**, 117, 1951.

A calculation analogous to that suggested by C.S. Smith for polygonal crystals [8]) may be employed for the flat crystals of Ferroxdure. Consider *fig. 6*, which shows a cross-section of a thin, flat crystal, of thickness δ and diameter D, making an angle Θ with the neighbouring mutually aligned crystals, of thickness d. If $\sigma_{\rm dt}$ be the boundary energy per unit area between the crystals meeting at angle Θ, and $\sigma_{\rm tt}$ that between the aligned crystals, the odd crystal will be able to grow provided that this growth reduces the overall boundary energy, that is if

$$2\,D\,\sigma_{\rm dt} < \frac{D\sin\Theta}{d}\,\frac{\delta}{\sin\Theta}\,\sigma_{\rm tt}.$$

It has been assumed in this calculation that angle Θ is not small, and that $D \gg \delta$.

From the above formula, then:

$$\frac{\delta}{d} > \frac{2\sigma_{\rm dt}}{\sigma_{\rm tt}}.$$

Now the boundary energy $\sigma_{\rm dt}$ increases very sharply with the deviation from the general orientation; in most cases, then, the incorrectly aligned crystal will be eliminated.

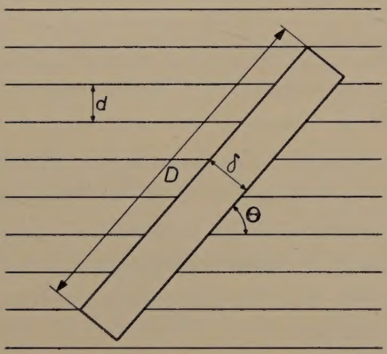


Fig. 6. Particle whose orientation deviates from that of the adjacent, mutually aligned crystals. As a general rule, such an oddly oriented crystal will be eliminated from the system.

Ferroxdure II and III

It will be seen from the above that any improvement in the anisotropy, that is, any increase in the remanence, is invariably associated with grain growth and therefore incompatible with a high coercivity (see table I). Accordingly, the theoretical maximum $(BH)_{\rm max}$ value has so far proved unattainable in practice, owing to the fact that when the remanence is sufficiently high, $\mu_0 {}_JH_{\rm c}$ $[{}_JH_{\rm c}]$ falls below the value corresponding to $B_{\rm r}/2$. Nevertheless, the $(BH)_{\rm max}$ values obtained by the method described above are considerably higher than those of isotropic material. It is possible to produce anisotropic materials with demagnetization curves specially suited to particular ranges of applications. At present, two types of anisotropic Ferroxdure magnet material are in production (see *fig. 7*).

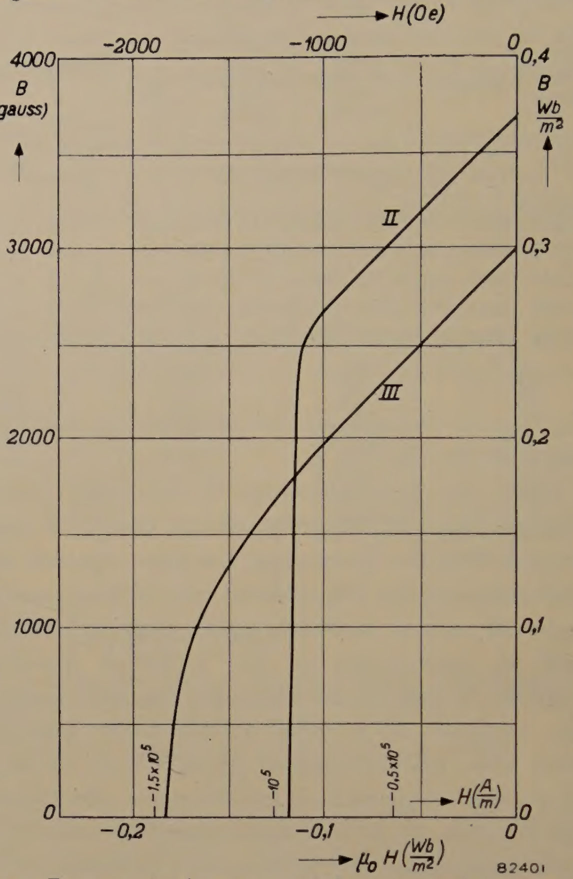


Fig. 7. Demagnetization curves of the anisotropic permanent magnet materials Ferroxdure II and Ferroxdure III.

One of them, Ferroxdure II, is intended for applications involving a fixed air-gap, in which high induction is the most important requirement, e.g. for loudspeakers. Its magnetic properties are as follows:

$B_{\rm r} = 0.35\text{-}0.40$ Wb/m² [3500-4000 gauss]
$\mu_0 {}_JH_{\rm c} = 0.11\text{-}0.15$ Wb/m² $[{}_JH_{\rm c} = 1100\text{-}1500$ oersted]
$(BH)_{\rm max} = 19\,000\text{-}24\,000$ J/m³ [$2.4\text{-}3.0 \times 10^6$ gauss oersted]

The other, Ferroxdure III, is intended for use in a strong demagnetizing field, or in circumstances where the working point may vary, as in magnetic clutches, dynamo magnets and yoke magnets. The magnetic properties of this material are:

$B_{\rm r} = 0.28\text{-}0.32$ Wb/m² [2800-3200 gauss]
$\mu_0 {}_JH_{\rm c} = 0.19\text{-}0.25$ Wb/m² $[{}_JH_{\rm c} = 1900\text{-}2500$ oersted]
$(BH)_{\rm max} = 16\,000\text{-}18\,500$ J/m³ [$2.0\text{-}2.3 \times 10^6$ gauss oersted].

---

[8]) C. S. Smith, Trans. Am. Inst. Mining Met. Engrs. **175**, 15, 1948.

It can be seen that in the case of Ferroxdure II, $\mu_0 {}_JH_c < B_r/2$ $[{}_JH_c < B_r/2]$, whereas for Ferroxdure III, $\mu_0 {}_JH_c > B_r/2$ $[{}_JH_c > B_r/2)$.

For the average properties of Ferroxdure types I, II and III at room temperature, see *table II.*

Table II. Average values of certain physical properties of Ferroxdure I, Ferroxdure II and Ferroxdure III at room temperature.

Magnetic moment per unit mass		8.75×10^{-5} Wb m/kg [70 gauss cm^3/g]			
Density		4800-4900 kg/m^3 [4.8-4.9 g/cm^3]			
Density of ideal single crystal		5300 kg/m^3 [5.3 g/cm^3]			
Electrical resistivity		$> 10^6$ Ωm [10^8 Ωcm]			
Temperature coefficient of induction		—0.2 % per °C			
		Ferroxdure			
		I	II	III	
Remanent induction		0.205	0.37	0.30	Wb/m^2
		2050	3700	3000	gauss
Coercivity	$\mu_0 {}_JH_c$	0.26	0.12	0.21	Wb/m^2
	${}_JH_c$	2600	1200	2100	oersted
Coercivity	$\mu_0 {}_BH_c$	0.15	0.12	0.20	Wb/m^2
	${}_BH_c$	1500	1200	2000	oersted
$(BH)_{max}$		7200	22 000	16 750	J/m^3
		$0.9 \cdot 10^6$	$2.8 \cdot 10^6$	$2.1 \cdot 10^6$	gauss oersted
Working point	B	0.11	0.26	0.17	Wb/m^2
		1100	2600	1700	gauss
	$\mu_0 H$	0.08	0.10	0.125	Wb/m^2
	H	800	1000	1250	oersted
Curie point		450 °C			

A high temperature coefficient of induction for points on the demagnetization curve is still an inconvenient feature of the anisotropic materials Ferroxdure II and Ferroxdure III.

A magnetizing force equivalent to $\mu_0 H = 1.40$ Wb/m^2 $[H = 14\,000]$ oersted is required to produce effective saturation in Ferroxdure I, whereas Ferroxdure II and Ferroxdure III can be saturated by a much weaker field.

In small-scale production, magnetic properties appreciably higher than those specified above can be attained. For example, a $(BH)_{max}$ of about 26 000 J/m^3 [3.3 × 10^6 gauss oersted] was obtained on each type of demagnetization curve, Ferroxdure II and Ferroxdure III.

The best result obtained so far is a $(BH)_{max}$ of 28 000 J/m^3 [3.5 × 10^6 gauss oersted] [9]); it will be seen from *fig. 8*, which shows the demagnetisation curve of the material concerned, that this is Ferroxdure II.

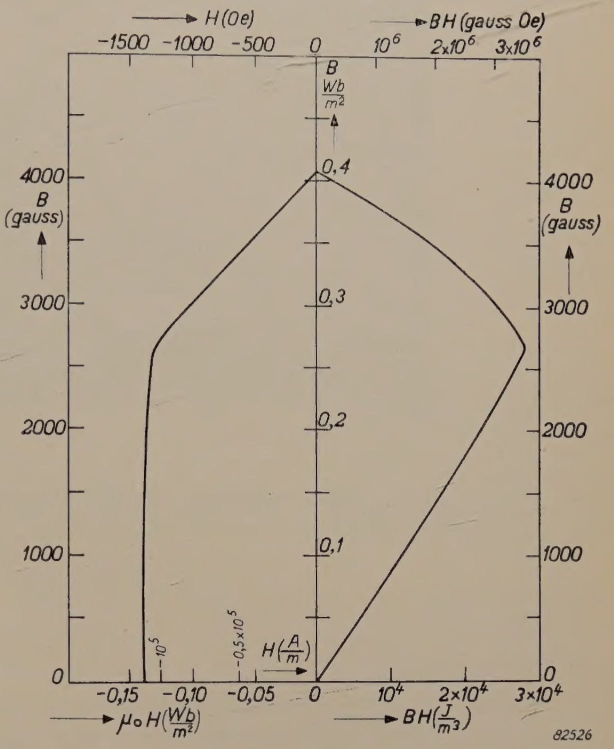
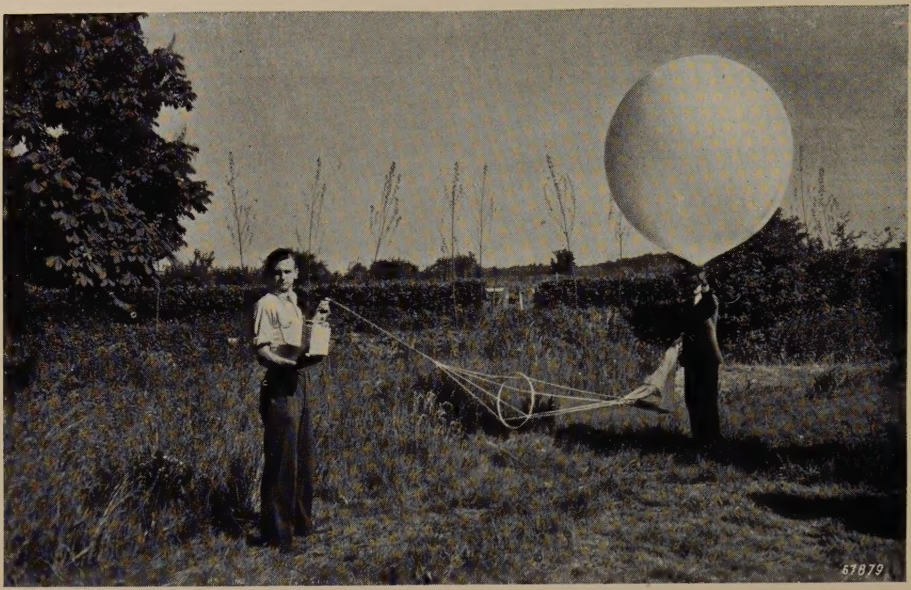


Fig. 8. Demagnetization curve of a particular sample of Ferroxdure II the $(BH)_{max}$ value of which is 28 000 J/m^3 [3.5 × 10^6 gauss oersted].

Summary. Owing to the low remanence of Ferroxdure I, the $(BH)_{max}$ value of this material is likewise low. The preferred directions of the particles constituting Ferroxdure I, an isotropic material, are distributed at random; hence the remanence of this material is no more than half its saturation value. However, by virtue of the distinct magnetic anistropy of the constituent crystals, the remanence of such materials can be increased considerably by press forming the powder in a magnetic field before sintering it. Before being so pressed, the Ferroxdure powder is ground until the diameter of the individual grains is less than the critical value for the formation of Bloch walls. Pressing in a magnetic field aligns the hexagonal axes of the grains more or less parallel to the direction of the lines of force, thus producing in this direction $(BH)_{max}$ values 3 to 4 times as high as that of the isotropic material. It is found that grain growth taking place during the sintering process enhances the magnetic anisotropy of the material considerably, although the decrease in coercivity caused by this growth limits the possible $(BH)_{max}$ value to some extent. Two anisotropic materials, Ferroxdure II and Ferroxdure III, each having its own particular range of applications, are described.

---

[9]) F. G. Brockman and W. G. Steneck, Philips tech. Rev. **16**, 79-87, 1954/55 (No. 3).



A RADIO SONDE FOR METEOROLOGICAL OBSERVATIONS

by A. HAUER *) and M. van TOL.

621.396.91

Radio sondes, as nowadays commonly employed by meteorological services, must satisfy fairly stringent requirements as regards accuracy, and must at the same time be as light and inexpensive as possible. The sonde developed by Philips in collaboration with the Royal Dutch Meteorological Institute and described in this article embodies several new features of importance in the fulfilment of these requirements.

Others, apart from the writers, who assisted in the development of this sonde were Messrs. R. J. Ritsma and H. J. A. Vesseur of the Royal Dutch Meteorological Institute, Mr. J. L. M. Reijnders of the Philips Measuring Equipment Development Dept., and Messrs. H. van Suchtelen and D. J. H. Admiraal of the Philips Research Laboratories in Eindhoven.

The radio sonde

The study of physical conditions in the upper-air, i.e. in that part of the atmosphere not within direct range of measuring instruments on the ground, is an essential branch of modern meteorology. Knowledge so acquired is particularly valuable in the preparation of weather forecasts.

The principal quantities to be measured are the temperature and the humidity as a function of altitude, but accurate pressure measurements are also required as a means of determining the altitude. Observations from aircraft enable us to record these data at altitudes up to about 10 km, but are very expensive and can be carried out only in the vicinity of an aerodrome; hence another method, i.e. balloon-borne measuring apparatus is preferred. Balloons have already been used for many years by meteorologists as a means of measuring wind velocity, but other information concerning the physical condition of the atmosphere can also be obtained from an ascent, by attaching recording equipment for pressure, temperature and humidity below the balloon. The latter, filled with hydrogen, climbs at a constant rate of about 5 m/sec., and bursts on reaching an altitude of 15-25 km. A parachute then opens to lower the recording equipment (sonde) slowly to the ground.

It is necessary to search for, and recover the sonde when once it has reached the ground, a task which may often present some difficulty, particularly in sparsely populated areas. Other serious disadvantages of this method are that the results of the observations cannot be ascertained until, and are never known unless, the sonde is found; hence a period of hours, or even of days may well elapse between the recording and the actual reading of the data.

---

*) Royal Dutch Meteorological Institute, De Bilt, Holland.

However, in about the year 1930, a method was evolved whereby the results of the measurements can be ascertained whilst still being gathered, that is, whilst the sonde is still airborne.

This method is that of the "radio sonde", now employed by most meteorological services. In principle, a radio sonde comprises two individual units. One of these, the measuring unit, contains instruments sensitive to temperature, humidity and pressure, and a circuit whose function is to pass the information to the other unit, that is, the transmitter. The latter is a small short-wave transmitter; the signal proceeding from this transmitter and received by equipment on the ground contains the results of the measurements in code form.

The radio sonde as a whole must satisfy one or two fundamental requirements, the most important of which is that it shall be sufficiently accurate (the accuracy considered ideal by meteorologists has not yet been attained by any radio sonde). It is likewise very important that the sonde be as light as possible, so that it will reach the highest possible altitude; this severely restricts circuit design, not least owing to the consequent limitation of battery capacity (the batteries make up a considerable proportion of the total weight). Hence the available power supply is very small.

To be suitable for regular, and at the same time economical use, the radio sonde must be inexpensive. In the Netherlands, the probability that it will not be recovered after an ascent is about 50 %, and even if recovered it may be seriously damaged (despite the parachute). In view of the latter possibility it is of course desirable that the radio sonde be mechanically robust.

In collaboration with the Royal Dutch Meteorological Institute at De Bilt, Philips have developed a radio sonde which embodies a number of unconventional features. A description will now be given, dealing with the items in the following order; the measuring unit, modulation of the transmitter by the measuring elements, the transmitter itself, and finally the results of several trial ascents by the new sonde.

Temperature, pressure and humidity measurements

To satisfy fully all the requirements of the meteorologists, a radio sonde would have to be capable of effecting extremely accurate measurements. In practice, however, the accuracy is limited to what is technically and economically practicable. Accordingly, the following requirements were taken as a starting point for the project here considered. Firstly, temperature measurements are to be accurate to within 0.5 °C, that is, in effect, to within 0.5 %, since the temperature in the particular regions of the atmosphere where the measurements are carried out varies on an average between 20° and —70 °C. Secondly, a similar accuracy is required in the measurement of pressure. The latter varies between 1050 and 70 millibars and the requirement is that it be measured accurately to within about 5 millibars, i.e. to within 0.5%. Thirdly, possible errors in the measurement of the relative humidity, which is expressed as a percentage of the saturation value, must not exceed 5—10 %.

The accuracy of such measurements depends firstly upon the particular measuring element in the sonde, secondly upon the manner in which the data are transmitted to the receiver on the ground, and thirdly upon the measuring instruments on the ground. In this article we shall consider only the first two factors.

Most radio sondes include several moving parts, the purpose of which is to convey the readings of the instruments measuring temperature, pressure and humidity mechanically to the transmitter, e.g. through a variable capacitance, resistance or self-inductance. Many also include a switching mechanism to connect the output of each measuring element in turn to the transmitter. This mechanism may be a rotary switch turned by a small windmill (which, in turn, is driven by the downward flow of air relative to the ascending sonde), a clock, or in some cases a small electric motor. Experience has shown that the reliability of sondes is often impaired by such mechanical parts.

In the Philips radio sonde the idea of switching was abandoned; measurements of the three quantities are transmitted simultaneously as three audio-frequency signals of variable frequency on a single carrier. This method necessitates the use of a separate audio-frequency oscillator for each of the quantities measured; altogether, then, three such oscillators are required, whereas if the signals were transmitted consecutively one oscillator would be sufficient. However, the switching mechanism has now been dispensed with, and by using suitable sub-miniature valves for the oscillators, the use of three oscillators has not increased either the total weight or the price of the radio sonde compared with earlier types. The oscillators will be considered presently. As we shall now explain, the number of moving parts in the sonde has been reduced to a minimum.

The measurement of temperature

The temperature measurement can be accomplished without the aid of any moving parts whatever by means of a temperature-sensitive element capable of acting direct upon the oscillator frequency, e.g. a resistor sensitive to temperature. This resistor must be made of a material having a temperature coefficient of resistance high enough to ensure the desired response. Such a material is to be found in the N.T.C. resistor, or thermistor [1]), that is, a semiconductor having a high negative temperature coefficient of resistance. It is found that a single N.T.C. resistor is capable of changing the frequency of an *R-C* oscillator by more than a factor of 2 for a decrease in temperature from 20° to —70 °C. For this purpose a straight-wire resistor was used whose resistance varies from about 30 kΩ at room temperature to about 500 kΩ at —70 °C. The change in the oscillator frequency per 0.5 °C (the desired accuracy) is then about 0.5 %, which can easily be measured.

The thermometer (see *fig. 1*) is a straight-wire N.T.C. resistor about 10 mm long, which is introduced direct into the frequency-controlling network of an audio-frequency oscillator (*R-C* oscillator). One advantage of this resistor over the bimetallic strip thermometer often used in other sondes is that it is very much quicker in response: it was found during trial ascents that the N.T.C. resistor responds quite noticeably to temperature fluctuations not detected by a bimetallic thermometer. Again, the "radiation error", which may be very appreciable in the case of a bimetallic strip (up to 10 °C), is very much smaller when a thin wire thermometer is employed. Hence the thermometer wire can be mounted quite freely about 10 cm outside the housing of the sonde, without any radiation screen or other cover.

The measurement of pressure

Metal aneroid barometers have been used as pressure-meters in almost all the radio sondes produced hitherto. However, the lag in response of such a barometer is very liable to introduce systematic error into the pressure measurement; moreover, the instrument itself is difficult to manufacture. For these reasons, and also to dispense with moving parts, another method based on the principle of the hypsometer was adopted, at the suggestion of the Meteorological Institute.

A hypsometer is a vessel, open to the atmosphere, containing a liquid maintained by some means at boiling-point; the boiling -point of the liquid depends upon the pressure of the surrounding atmosphere, that is, it decreases with this pressure. If the precise relationship between pressure and boiling-point is known, then, the one can be determined by measuring the other (e.g. by means of an N.T.C. resistor immersed in the liquid). The hypsometer is very reliable as a pressure-meter provided that the liquid used in it is sufficiently pure, since for pure liquids the relationship between boiling-point and pressure is very well established.

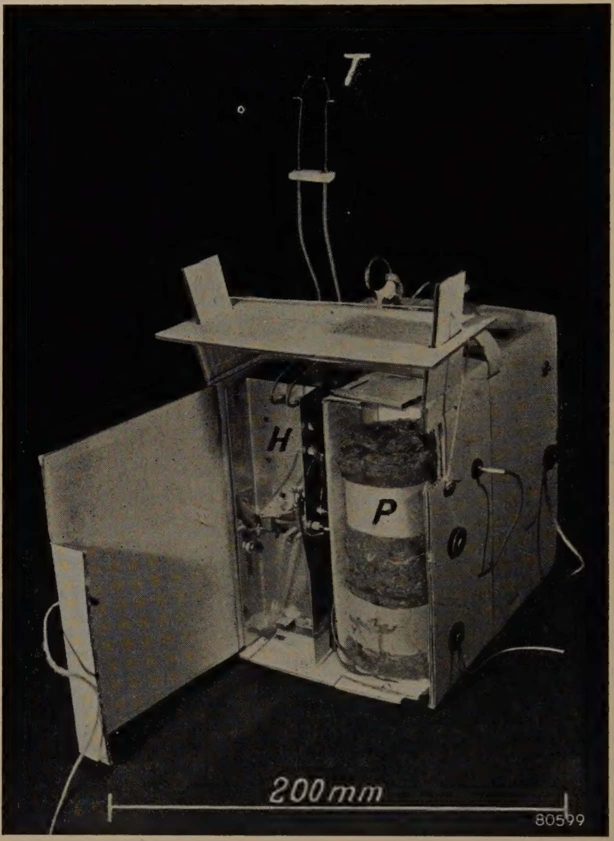


Fig. 1. The Philips radio sonde. The entire instrument is housed in a cardboard case, 165 × 120 × 155 mm. The temperature-measuring element *T* is a straight-wire N.T.C. resistor 10 mm long, mounted outside the sonde; the pressure and humidity measuring-elements, *P* and *H* respectively, are partly visible at the open side of the sonde. Altogether, the sonde weighs 500 grams; it is suspended from a hydrogen-filled balloon by a line about 10 m long, 5.5 m of which is covered with conducting sleeving to act as an aerial.

To maintain the liquid at boiling-point, at which some of it evaporates, a certain amount of energy is required. This would of course be a big disadvantage if batteries were the only source of energy available, since these would increase the weight of the sonde. However, if we have a liquid whose boiling-point will remain below the temperature of the surrounding atmosphere throughout the ascent, the energy to keep the liquid boiling will be supplied by the atmosphere itself. This is in fact practicable since although the temperature of the atmosphere drops as the radio sonde ascends, the accompanying decrease in pressure causes a similar drop in the boiling-point of the liquid. Curves *1* and *2* in *fig. 2* show the relationship between decreasing pressure and decreasing temperature in the atmosphere (these curves are based on the summer and winter averages, respectively, of observations carried out at various times and places). Accordingly, we must choose a hypsometer-liquid whose pressure/boiling-point curve is to the left of curve *2* in fig. 2. At the same time the two curves should not be too far apart or the liquid will evaporate too quickly and it will be necessary either to insulate the hypsometer heavily, or to fill it with a very large amount of liquid. Freon, the vapour pressure of which is indicated by curve *3* in fig. 2, is a liquid which satisfies these conditions very well indeed.

---

[1]) E. J. W. Verwey, P. W. Haaijman and F.C. Romeyn, Semi-conductors with large negative temperature coefficient of resistance, Philips tech. Rev. 9, 239-248, 1947/1948.

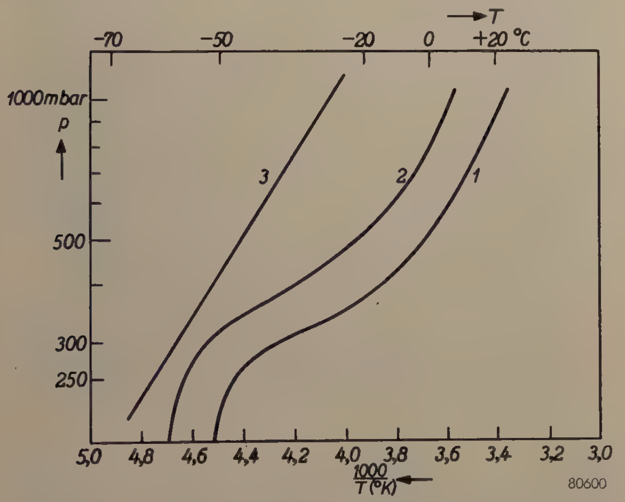


Fig. 2. Variation of temperature with pressure in the "average" atmosphere during summer (curve *1*) and winter (curve *2*) Line *3* represents the vapour-pressure curve of freon. The pressure of the atmosphere *p* is plotted on a logarithmic scale against the reciprocal of the absolute temperature *T*, so that curve *3* is then nearly a straight line.

There is a limited region of the atmosphere whose average temperature (in winter) is very close to the boiling-point of freon, so that on cold days the temperature of the air in this region may drop below the boiling-point of the hypsometer liquid. Apparently, then, freon fails to satisfy the condition governing spontaneous boiling, but despite this the liquid is prevented from "going off the boil" by a second source of heat within itself. This may be explained as follows. During the ascent of the sonde the entire mass of liquid, the temperature of which cannot of course be any higher then is consistent with the continually decreasing boiling-point, must dissipate more and more of its heat content. This heat helps to promote evaporation and the power thus made available will be all the greater, the more rapid the decline of the boiling-point, that is, the higher the rate of ascent of the sonde. Provided that the rate of heat transmission through the insulation of the boiling-vessel to the relatively cold surroundings is less than the rate at which energy from the heat of the liquid becomes available, a certain surplus of heat will always be available to keep the liquid boiling.

It is found that a relatively small quantity of freon (10 cm^3), insulated with 1 cm thickness of cotton wool, is sufficient to supply the hypsometer during an entire ascent. The vessel used to contain the liquid is an ordinary radio-valve bulb; the N.T.C. resistor is secured to two lead-in pins through the bulb in such a way as to be completely immersed in the liquid (*fig. 3*). The resistor, like the air-temperature measuring element, is part of the frequency-controlling network of an *R-C* oscillator. The outer air has access to the liquid through a narrow tube (exhaust stem) at the top of the bulb; a completely open vessel would be unsuitable since, apart from the chance of spilling, it would allow the freon to evaporate too quickly and so perhaps assume a temperature below the boiling-point. On the other hand, access to the surrounding atmosphere should not be too severely restricted, in view of the possibility of the liquid overheating.

Tests have shown that a pressure variation of

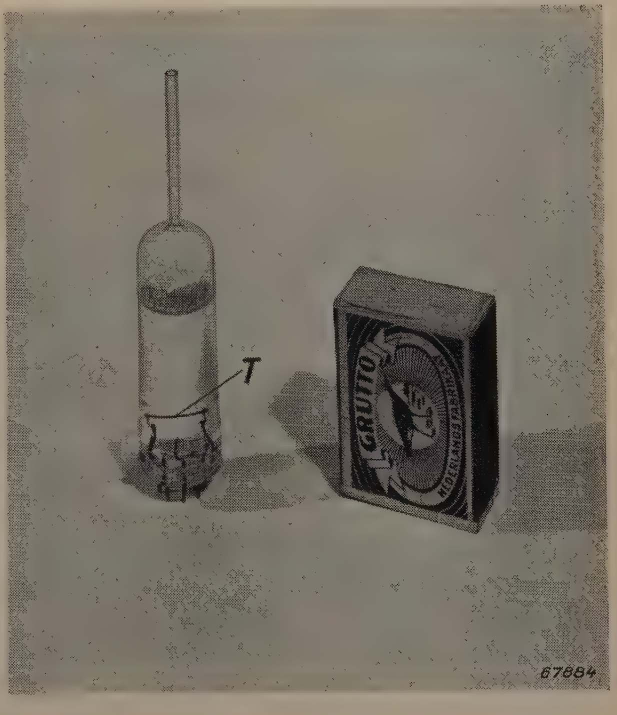


Fig. 3. Hypsometer. This is made from a radio-valve bulb containing freon, the boiling-point of which is measured by means of an N.T.C. resistor *T* secured to two lead-in pins in the bulb and completely immersed in the liquid; the latter has access to the outer air through a tube (the exhaust stem) at the top of the bulb.

5 millibars corresponds to a variation of the boiling-point such as to produce a relative variation of 0.2 % or more in the oscillator frequency.

The measurement of humidity

In the Philips radio sonde, as indeed in several others, a piece of gold-beater's skin is used as a means of measuring the relative humidity of the atmosphere. This skin is an animal tissue, the chief constituent of which is keratin, and like human hair, expands and contracts with changes in the humidity of the surrounding atmosphere. Unfortunately this material is affected by hysteresis when it passes from a humid to a dry state, and then back to the humid state. For this reason the electrical resistance of a solution of lithium chloride is employed instead of gold-beater's skin as a measure of humidity in sondes of one American make. This is a very sensitive method, but one which requires alternating current to make it fully effective and moreover, involves the use of rather complex measuring equipment; hence gold-beater's skin is preferred in the radio sonde described here, despite the above-mentioned defect.

The variation in the length of this skin affects the width of the air-gap in the magnetic circuit of a coil (*fig. 4*) and is thus converted into a variation of the self-inductance of this coil. The coil is part of an audio-frequency *L-C* oscillator, the frequency of which varies by approximately a factor of 2 for a variation in relative humidity from 10 to 100 %.

The audio-frequency oscillators

As we have already seen, mechanical switching is dispensed with in the Philips radio sonde by virtue of the fact that the three audio-frequency signals corresponding to the three quantities to be measured are modulated on one carrier wave and transmitted simultaneously. These signals reach the ground station receiver continuously, and the response time of the recording equipment connected to this receiver does not have to be very small since only gradual variations take place in any of the quantities measured.

Accordingly, there is an audio-frequency oscillator corresponding to each measuring element in the sonde. Since the sensitive elements in the temperature and pressure measuring devices are resistors, the obvious course was taken in adopting *R-C* circuits for the two associated oscillators. As already explained, the oscillator coupled to the hygrometer is of the *L-C* type. The frequencies of the three oscillators are adequately spaced; hence the three signals are readily separated in the receiver.

The frequency bands are as follows:

1— 3 kc/s for pressure;
4— 8 kc/s for temperature;
12—25 kc/s for relative humidity.

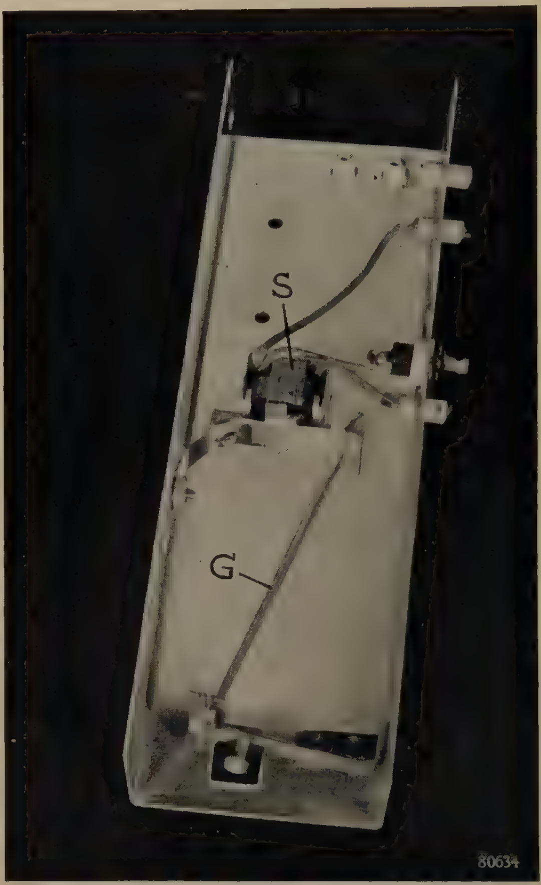


Fig. 4. Hygrometer. The variation in length of a piece of gold-beater's skin *G* is converted into a variation in self-inductance by virtue of the fact that the skin controls the width of the air-gap in the magnetic circuit of coil *S*.

The *R-C* oscillators must satisfy certain requirements:

1) The input power must be as low as possible (with a view to the weight of the batteries).
2) The number of components should be as small as possible (with a view to the price and total weight of the sonde).
3) All the components of the different networks (apart from the N.T.C. resistors, of course) should remain as far as possible unaffected by temperature and pressure variations.
4) The circuit should be insensitive to battery-voltage variations.

To satisfy requirement 1), sub-miniature valves are employed. These are very small valves designed to operate with batteries; the type used in the oscillators is the DL 67, which has a diameter of 7.5 mm and a length of 35 mm. Valves of the same type are used in hearing-aids, etc., and were originally developed for military purposes, e.g. for proximity fuses. *Fig. 5* shows the DL 67, together with a valve of the type DL 41 also employed in the sonde. The former are wired direct into the circuit.

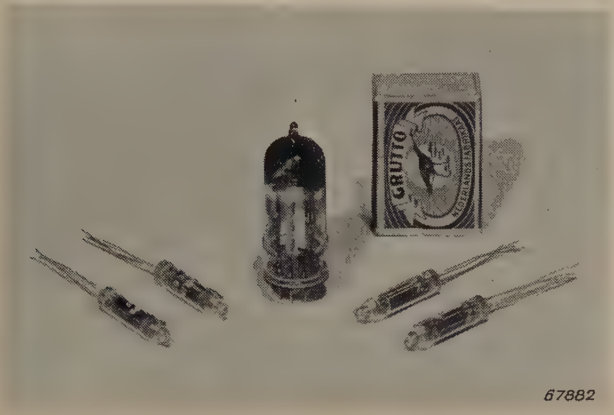


Fig. 5. The five valves used in the radio sonde. The four small valves are battery-operated sub-miniatures type DL 67 (length 35 mm, diameter 7.5 mm), three of which are employed in the three oscillators and the fourth as a reactance valve. Centre: the battery-operated valve type DL 41 (dimensions 20 × 55 mm) used as the transmitting valve of the radio sonde.

In as far as they relate to the *L-C* oscillator, the other requirements can be satisfied merely with the aid of standard components (of the smallest possible size), but the *R-C* oscillators present a more difficult problem. The circuit diagram of an *R-C* oscillator is shown in *fig. 6*, from which it will be seen that a triple *R-C* network is employed as feedback between anode and control grid (this is the minimum number of components required to produce the phase-inversion necessary for oscillation). Small carbon resistors and ceramic or miniature mica capacitors are used in the networks.

Temperature variations affect the circuit mainly through the carbon resistors (and the batteries), the capacitances being virtually independent of the temperature. However, the effect of such variations can be minimized by insulating the circuit as thoroughly as possible from the surrounding air, so as to prevent the transfer of heat.

There is a direct connection between the above and the fourth requirement, which may be explained in the following manner. During the ascent of the sonde there is a slight decrease in the battery voltages, partly owing to the low capacity of the batteries, and partly to the fact that these voltages also depend upon the temperature. To fully appreciate the deleterious effect of this decrease, we must re-examine the diagram shown in fig. 6.

If the circuit is to oscillate, the attenuation caused by the network of capacitors and resistors must be less than the amplification effected by the valve; at the same time, an unlimited increase in the amplitude of the oscillations must be avoided. Normally, then, such an oscillator circuit will contain either an incandescent lamp or an N.T.C. resistor, so positioned that the positive feedback of alternating current in the circuit decreases with the amplitude of the oscillations. However, amplitude limitation by this method cannot be employed in our circuit owing to the fact that the A.C. power available is very low. Hence we make use of a certain "natural" restriction of amplitude arising from the fact that the grid of the valve is conducting during a small part of each cycle. As a result of this, the control grid acquires a negative bias (the grid capacitor is charged), which reduces the mutual conductance of the valve. As the amplitude of the oscillations increases, the voltage on the control grid becomes more and more negative, until the mutual conductance is so reduced that the amplification by the valve is only just sufficient to composate for the attenuation by the network. The amplitude of the oscillations then remains constant.

However, during the time that the grid is conducting, a new conducting link is formed between the grid and the cathode (effective for A.C. as well as for D.C.), so that in effect an extra resistance R_s is introduced in parallel with the terminating impedance R of the network (*fig. 7*); hence a slight change takes place in the frequency of the oscillator, which is governed by the capacitance and the resistances of the network.

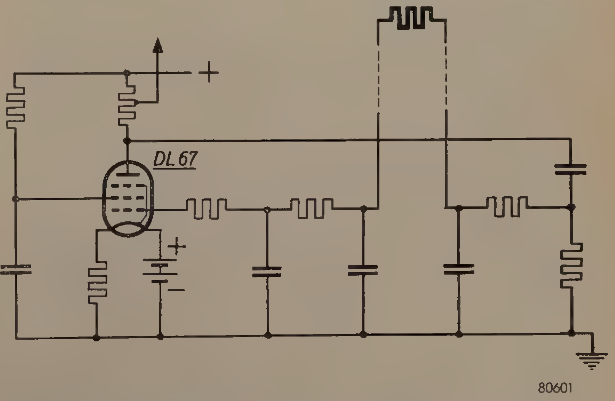


Fig. 6. Circuit diagram of *R-C* oscillator, two of which are employed in the radio sonde. The measuring resistor is drawn in bold lines.

Although insignificant in itself, this frequency shift is governed by the battery voltage, since the amount of the extra resistance depends upon this voltage. The reason for this as follows. Following a reduction of the battery voltage, and thus a reduction of the mutual conductance of the valve, a relatively smaller negative grid voltage is sufficient to bring about the further decrease in mutual conductance then required. Moreover, the grid current is also somewhat reduced during this process When the grid A.C. voltage is high, the grid current is, to a first approximation, proportional to it, so that R_s is then constant. This proportionality does not exist however at the low grid voltages which occur in the present circuit. In the case here considered, then, R_s depends upon the grid current, that is, upon the battery voltage.

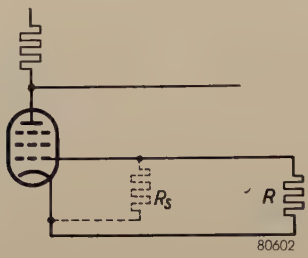


Fig. 7. Equivalent circuit of the grid circuit of the oscillator valve in an *R-C* oscillator, showing the apparent resistance R_s resulting from the flow of grid current. *R* is the total terminating impedance of the filter.

To minimize this effect two measures were adopted. Firstly, a high-value resistor was included in the screen grid circuit to moderate the effect of battery-voltage variations upon the screen grid voltage, that is, upon the mutual conductance of the valve. Secondly, the control grid was given an initial negative bias such as to reduce the grid current as far as possible without causing too great a decrease in amplification. Since the heater voltage of the DL 67 is only 1.25 volts whereas the available battery voltage is 2.3 V, the required bias is readily extracted from the difference between the two by including a resistor in the cathode circuit.

As it happened, the battery effect persisted despite these two measures to prevent it. However, it was found that the frequency shift is also governed by the anode resistance, and since, amongst other things, the decrease in the anode-battery voltage causes a frequency shift opposed to that resulting from the decrease in the voltage of the heater battery, the entire effect can be practically eliminated by choosing a suitable anode resistance. Experiments have shown that in this way the frequency shift caused in each oscillator by a decrease of about 10% in the two battery voltages can be limited to less than 0.2%.

In practice the decrease in voltage during an ascent is invariably less than 10% [2]) and, as will be seen from the accuracy requirements already defined for the pressure, temperature and humidity, the frequency shift is then sufficiently small.

The transmitter

It was at first intended to employ 1 metre as the wavelength of the high-frequency carrier to be modulated by the three audio-frequency signals, The position of the sonde could then be determined at regular intervals during the ascent by means of a radio-theodolite in order to measure wind-velocities, as was, in fact, originally the principal object of such balloon ascents. However, since the bearings so obtained become rather inaccurate when once the sonde has travelled any considerable distance from the ground-station (>50 km), and also owing to the difficulty of designing a simple, one-valve high-frequency oscillator suitable for modulation at such a short wavelength, the idea of position-finding was abandoned for the time being and a wavelength of 11 metres (28 Mc/s) was adopted for the carrier wave (this wavelength is employed in several other radio sondes). The battery valve type *DL* 41 (see fig. 5) is used for the transmitter, since it supplies sufficient power for our purpose (about 100 mW).

Fig. 8 shows the circuit diagram of the high-frequency oscillator; the aerial is a conducting sleeve one half-wavelength in length round the wire on which the sonde is suspended from the balloon. In principle, either amplitude modulation or frequency modulation could be employed; the latter invariably necessitates the use of an intermediate stage, whereas, in theory at least, the former enables the modulating voltage to be applied direct to the control grid of the transmitter valve. However, in the case here considered the output power of the oscillators is insufficient for direct application, and since for this reason alone it is essential to include an intermediate stage, frequency modulation is preferred, by reason of the fact that the signal-to-noise ratio is very much more favourable. Accordingly, this method of modulation was adopted.

The three signals of the "pick-up" oscillators, separated by intervals between the frequency bands, are mixed in a potentiometer circuit (see fig. 8) and are applied to the grid of a "reactance" valve (again a DL 67), the principle of which has already been described in this Review [3]); the effect of such a valve is the same as that which would be produced by a capacitance or self-inductance (in this case a capacitance) in parallel with the high-frequency oscillator circuit, its value being governed by the negative voltage on the grid of the valve. This voltage, which varies at audio frequency with the frequencies of the three oscillator-signals, thus modulates the frequency of the oscillatory circuit.

---

[2]) In fact, the batteries are accumulators, specially designed for use in radio sondes, the voltage of which depends only very slightly upon temperature.

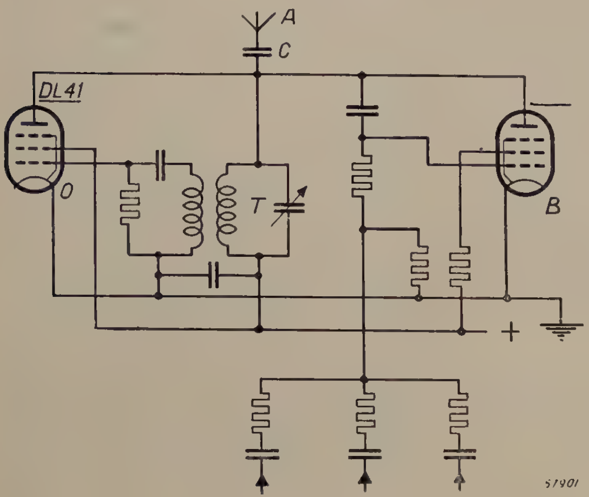


Fig. 8. Circuit diagram of the high-frequency oscillator containing the reactance valve. The three audio-frequency signals are mixed in a potentiometer and applied to the grid of the reactance valve *B*. The frequency of the H.F. oscillator *O* is governed by the voltage on this grid. The oscillator is connected across a capacitance *C* to the $\lambda/2$ aerial *A*. The frequency of the oscillator can be adjusted by means of trimmer *T* just before the ascent of the sonde.

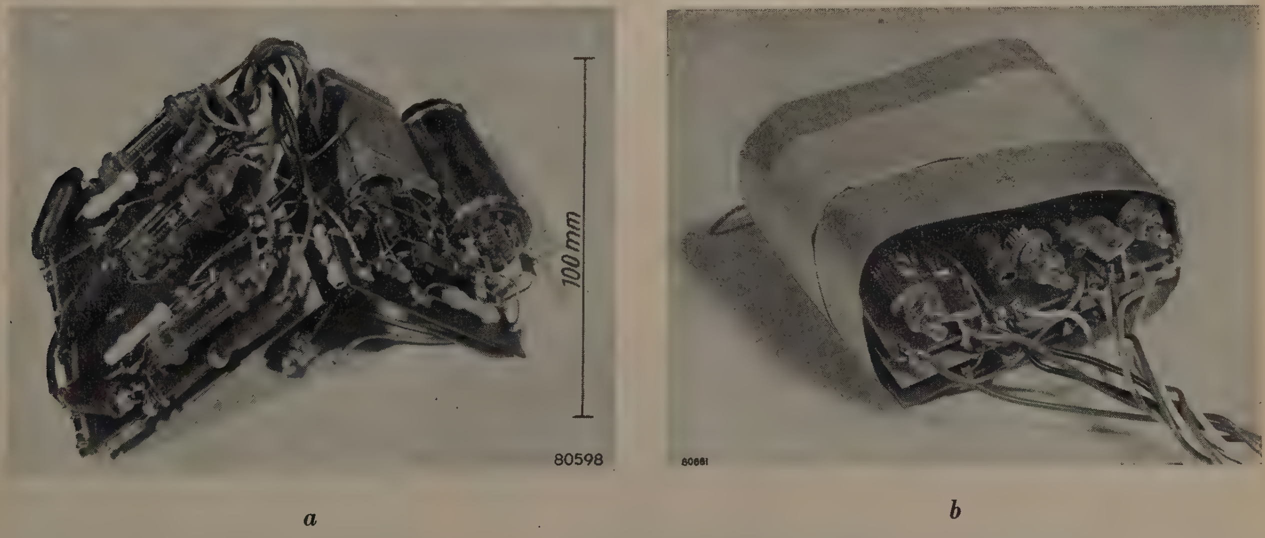
One factor affecting the accuracy of the meteorological data as received on the ground is the frequency sweep of the transmitter; the greater this sweep, the stronger the audio-frequency signal in the receiver. In this case, a maximum frequency sweep of 25 kc/s was adopted as being the most consistent with the accuracy requirements imposed and the possibilities of the circuit.

Constructional details of the sonde; features of the ground-station equipment

The high-frequency section is completely screened from the remainder of the sonde, to eliminate possible interference; the carrier-wave frequency can be changed slightly just before an ascent by means of a small trimming capacitor provided for the purpose.

The transmitter and the oscillators are together accommodated in a robust cardboard cover; this greatly increases the probability that they will be recovered intact. *Fig. 9* shows the interior structure of the sonde.

A brief description of the ground-station receiving equipment will suffice. In principle, this may comprise a frequency-modulated receiver with three audio-frequency bandpass filters, one for each of the quantities to be measured, and three frequency meters to determine the frequencies of the signals passed by the filters. Preferably, the meters used to measure the frequencies should be of the recording



a

b

Fig. 9. *a*) Electronic section of the radio sonde. The three oscillators are mounted on the left-hand panel and the transmitter on the right-hand panel. *b*) In the assembled sonde the two panels lie back to back and are enclosed in a cardboard cover. The measuring elements and batteries are connected to the panels by flexible leads.

---

[3] For exemple, see Th. J. Weyers, Frequency modulation, Philips tech. Rev. **8**, 42-50, 1946 (particularly page 47).

type. *Fig. 10* shows the receiving equipment at present installed in the Royal Meteorological Institute at De Bilt.

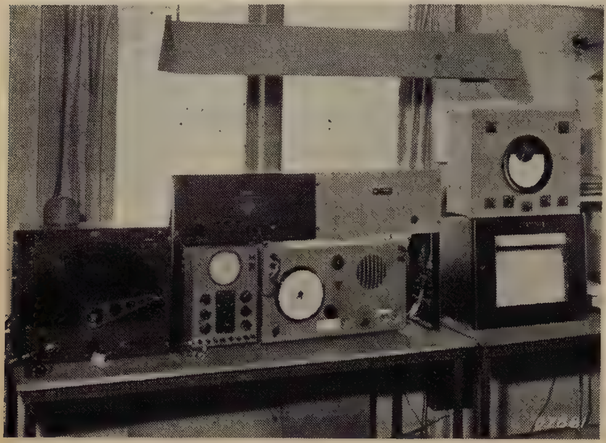


Fig. 10. Photograph of the ground-station receiving equipment as installed in the Royal Dutch Meteorological Institute. The F.M. receiver is seen at the centre, and the recording meter on the right. At the extreme left is the calibrating equipment which supplies a constant comparison-frequency. Calibration is preferably effected with the aid of a cathode-ray oscillograph.

Testing the radio sonde

To test the accuracy of the radio sonde described in this article, the Dutch Meteorological Institute carried out several "twin-ascent" trials, that is, they sent up two such sondes suspended from one balloon, and compared the temperature and pressure data transmitted by these instruments. It was found as a result of these trials that the pressure measurements of the two sondes agreed to within 5 millibars up to an altitude of 5000 metres, and that the differences at altitudes up to 10 000 metres did not exceed 10 millibars. The temperatures so measured agreed to within 0.5 °C.

Although the number of observations carried out during these trials is too small to justify any definite conclusions, the results show that the pressure measurement is at least as accurate, and the temperature measurement about twice as accurate, as those quoted in the literature for other sondes. Moreover, it is found that the results produced by the hypsometer at the temperatures occurring within its practical range of operation are entirely independent of the temperature: hence the pressure measurements of the sonde do not require correction to compensate for possible temperature fluctuations.

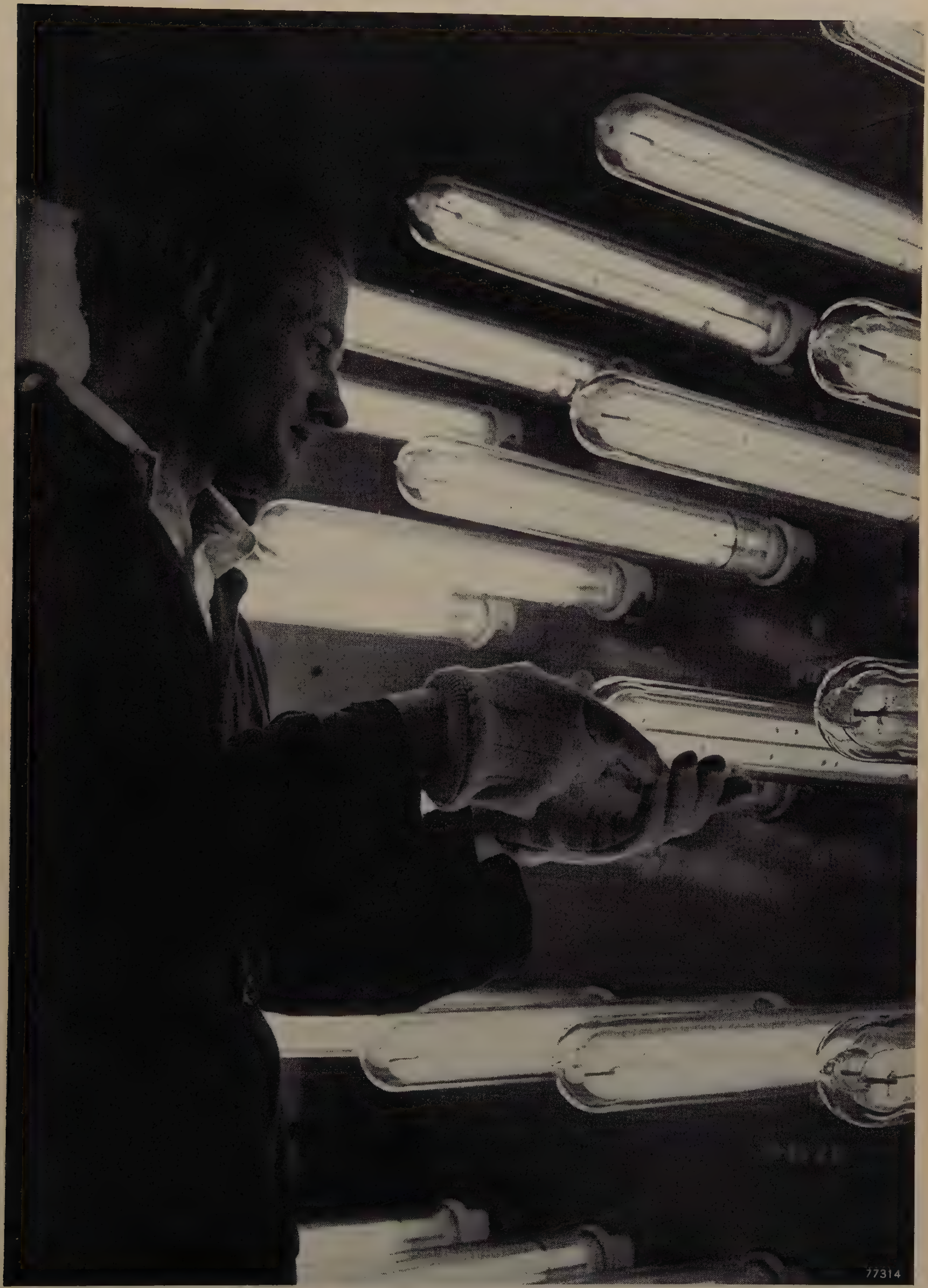
Again, it appeared from the twin-ascent trials that the new sonde is very convenient to handle, and that it is not easily damaged. Another important feature of the sonde revealed by these trials is its relative lightness; the combined weight of two of the new sondes is still less than that of one of the type hitherto employed.

A number of the new instruments were recovered after ascents, and there were very few instances of damage to the electronic equipment contained in them. In fact, one of the advantages of the new sonde is that the temperature and pressure calibrations are not affected by damage other than to the measuring resistors, since they are governed solely by the electrical properties of the sonde. This was verified by re-calibration of the sondes recovered.

Another point worth mentioning in connection with the calibration of the new sonde is that this can be accomplished without exposing the instrument as a whole to the effects of cold at the low calibrating temperatures. It is sufficient to plot the resistance v. temperature curves of the resistors, and the frequency curves of the oscillators as a function of the particular values of the variable measuring elements.

Summary. The instrument described as a "radio sonde" is used as a means of observing physical conditions in the atmosphere; in it, elements for measuring temperature, pressure and humidity are combined with a short-wave transmitter emitting a signal which contains the information gathered by the three elements. All this equipment is attached to a balloon, which is then sent aloft. Such a sonde must satisfy a number of stringent requirements, the most important of which are that it shall be accurate, light and inexpensive. The sonde described in this article differs from earlier types in that it transmits the three items of information simultaneously and so requires no switching mechanism. Each measuring element controls a small oscillator, the frequency of which is influenced by the particular variable to be measured. Pressure-variations are transmitted in the frequency band between 1 and 3 kc/s, temperature-variations in the band between 4 and 8 c/s, and variations in humidity in the band between 12 and 25 kc/s. To minimize the number of moving parts required in the sonde an N.T.C. resistor, included in the appropriate R-C oscillator circuit, is employed as a thermometer, and the pressure is ascertained from the boiling-point of freon, which is likewise measured by an N.T.C. resistor. The relative humidity is established with the aid of a piece of gold-beater's skin, variations in the length of which affect the self-inductance of a coil forming part of an *L-C* oscillator. The description of the R-C oscillators includes an account of how the frequencies of these oscillators are rendered independent of the battery voltage, which decreases to some extent during the ascent of the balloon. Sub-miniature valves are used in the oscillators; altogether, the sonde contains four of these valves. Frequency modulation is employed, the wavelength being 11 metres and the maximum frequency sweep 25 kc/s. The total output power of the transmitter valve (the battery-operated DL 41) is about 100 mW. Particulars of the mechanical design of the sonde and the principles of the ground-station receiving equipment are given in brief. Finally, the results of a number of practical tests on the new radio sonde are described.

SEASONING OF SODIUM LAMPS



The photograph shows a group of sodium lamps during the "seasoning" process — a stage in the manufacture during which the sodium is distributed in droplets along the discharge-tube, in order to ensure a uniform brightness in the finished lamp.

A PROFESSIONAL CINE PROJECTOR FOR 16 mm FILM

by J. J. KOTTE.

778.55.068.25

Originally 16 mm film was introduced to offer film-making facilities to the amateur at a reasonable cost. The advent of 8 mm film, however, was an even more important step in bringing amateur film making within the reach of the general public, and has now largely supplanted 16 mm film for amateur use. Nevertheless, 16 mm film survived, and acquired a range of applications peculiar to itself, including, for example, instructional and educational films. It is nowadays employed mainly by professionals, whose standards, as regards sound and picture quality as well as general reliability of equipment, are far more exacting than those of the amateur.

Most of the latest 16 mm projectors are simply more elaborate, and improved forms of the original equipment designed for amateurs. However, in the Philips projector, which has now been in production for some time, every link with the early amateur equipment has been severed; nevertheless, this projector is so simple to operate that it can be used even by persons not specially trained for the purpose.

For some time now, Philips have been producing in addition to standard 35 mm film projectors, a conveniently transportable 16 mm sub-standard film projector, type EL 5000, specially designed for use by professionals. *Fig. 1* is a photograph of this projector. The need for such a projector arises from the fact that 16 mm film has now largely passed from amateur, into professional hands [1]).

In view of this fact, special consideration has been given to such features as high luminous flux, to permit of projection on to a relatively large screen (up to 4 × 3 m), and robust construction, to enable the projector to be employed continuously for several hours per day, as well as to the overall quality of picture and sound. The total luminous flux falling upon the screen from the projector, using a 750 watt lamp and in the absence of film, is 500 lumens, which compares very favourably with the values [2]) usually rated as "acceptable" (200 lumens) and "very good" (275 lumens).

The means employed to secure this very high light flux will now be described, and in the course of this description reference will be made to the optical system, the shutter and the intermittent mechanism. Being in some respects novel in design, the last of these three items will be disdiscussed more fully than the others. Next, several details of the design, some of which are shared by Philips standard 35 mm projectors, will be considered, and, finally, a concise account of the sound system will be given.

The luminous flux

The theoretical maximum of luminous flux Φ_{th} that can be projected on to the screen is governed entirely by the area (S), the luminance (B_s) of the film gate, and the angle α subtended by the effective diameter of the objective at the film gate (*fig. 2*).

This flux is given by [3]):

$$\Phi_{\mathrm{th}} = \pi B_s \cdot S \sin^2 \tfrac{1}{2}\alpha .$$

According to a well-known optical principle, the luminance along a beam of light proceeding through an optical system devoid of absorption and reflection losses is constant. However, since losses caused by absorption and reflection occur in every optical system, B_s is invariably less than the luminance B of the light source. Similar losses occur in the objective. The shutter gives rise to another unavoidable loss, which will be examined more closely later on. Taking all these losses into account by means of a factor k, we find that the total luminous flux emitted by a projector is:

$$\Phi = k \cdot \pi BS \sin^2 \tfrac{1}{2}\alpha \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Since the area of the film gate is governed by the size of the film (being 0.75 cm² for 16 mm film), maximum light yield can be obtained only by increasing B, $\sin^2 \frac{1}{2}\alpha$ and k to the fullest extent.

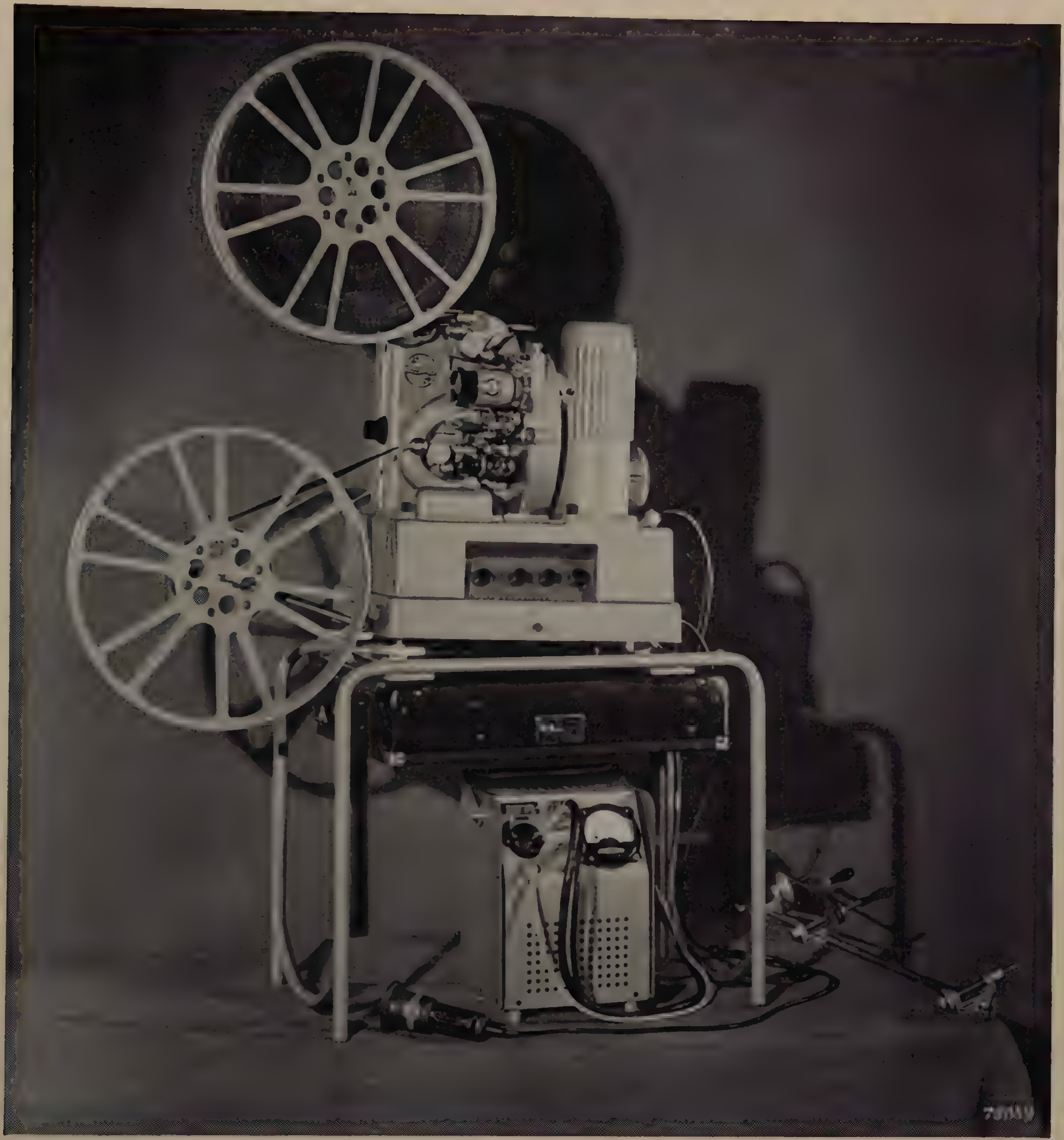
Formula (1) is not strictly valid unless the particular lighting system (in this case comprising the lamp and a condenser) is so designed that all rays completely fill the objective, irrespective of the point at which they pass through the film gate (*fig. 3a*). In principle, any projector lighting-equipment can be so designed, but in practice, using a light source of

---

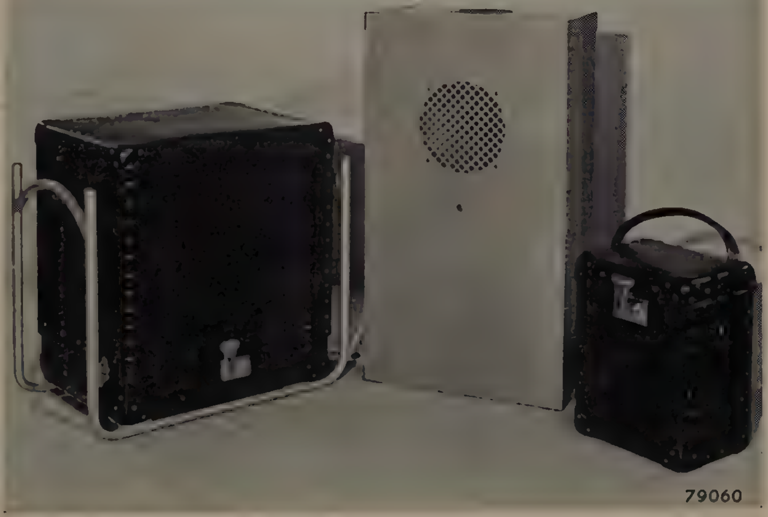
[1]) Many aspects of 16 mm film, including industrial applications. are discussed in detail in "Sixteen mm Sound Motion Pictures" by W. H. Offenhauser, published by Interscience Publishers, New York, 1949.

[2]) See page 352 and 453 of the book by Offenhauser referred to above.

[3]) See G. Heller, A film projection installation with water-cooled mercury lamps, Philips tech. Rev. 4, 2-8, 1939.



a



b

Fig. 1. *a*) The 16 mm sound film projector EL 5000 ready for use. It is seen mounted on a standard support which brings the top of the upper film spool approximately to head level. The transformer for converting the mains voltage to 110 Volts will be seen beneath the projector. On the left of this transformer is a hand microphone, and at the extreme right a manual re-winder.
b) The equipment packed for transport in three cases, that on the left containing the projector, that in the centre the two loudspeakers, associated cables, spools and re-winder, and that on the right the transformer.

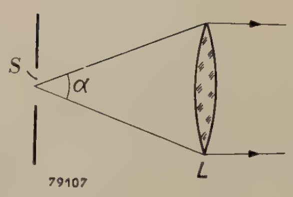


Fig. 2. The film gate (*S*) is uniformly illuminated by the lighting system (not shown). Since the projection screen is very remote from the projector the film gate is positioned for all practical purposes in the focal plane of the objective (*L*), here represented, for convenience, by a single lens.

moderate size and a reasonably inexpensive condenser, it is necessary to effect a compromise which satisfies the above condition only in respect of points in the central zone of the film gate. Rays passing through points outside this zone fail to fill the objective completely (fig. 3*b*) and so produce illumination which fades gradually towards the edges of the projection screen. Another reason for this fading is that rays arriving at a large angle to this axis are partly obscured by the edge of the objective (vignetting).

It is usual in cinematography to quote as the luminous flux the particular value corresponding to the illumination measured at the centre, assuming uniform illumination over the whole screen; this is in view of the fact that the measurement of the real luminous flux (which is always smaller) is a rather complex process.

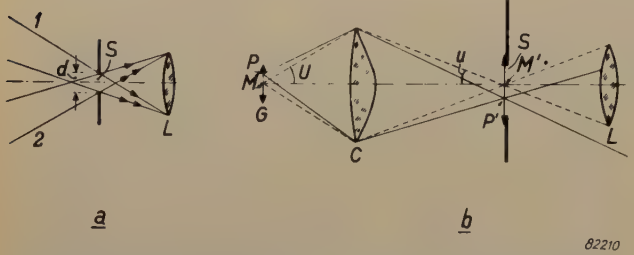


Fig. 3. *a*) To make full use of the objective it would be necessary to illuminate the film gate in such a way that even the light rays proceeding from the corners of this gate completely fill the objective. If *d* be the diagonal of the film gate, light rays *1* and *2* must still be present.
b) If the condenser *C* is employed to form an image of the light source *A* in the film gate, and the angle subtended at the centre *M'* (the image of *M*) of the film gate (*S*) by the condenser is the same as the angle subtended there by the objective (*L*), the beam passing through point *M'* will completely fill the objective. On the other hand, a beam passing through another point, *P'* (image of *P*), chosen at random in the film gate, will not fill the objective completely.

The light source

The most suitable light source for a transportable projector is the incandescent lamp, since this is relatively small and light in weight, and can be connected to the mains without any complex ballast. For various practical reasons, it was decided to employ a 750 Watt, 110 Volt lamp, the rated life of which is 25 hours. (The filament temperature, and hence the luminance level, may be increased, provided that a shorter lamp-life is considered acceptable). Quantity *B* in formula (1) is the average luminance of the area covered by the spiral filament). To procure the highest possible luminance *B*, a bi-plane filament is employed [4]) as seen in *fig. 4*, with a concave mirror mounted behind it. This mirror is actually formed on the inside of the lamp envelope, part of which has the form of a sphere. The cylindrical shape of the usual incandescent lamps employed in sub-standard film projection (fig. 4) necessitates the use of an external mirror, which, to protect it from damage, is usually silvered at the back. Hence the light reflected from such a mirror to the screen must pass six more air-glass interfaces than it would if similarly reflected from an internal mirror. Since a reflection loss of about 4% occurs at each boundary plane, the internal mirror is obviously more efficient. Other advantages of the internal mirror are that it is factory-adjusted, does not tarnish readily, and is renewed automatically whenever a lamp is replaced. Another, incidental advantage of the spherical lamp shape is that, indirectly, it improves the cooling; this may be explained as follows. The lamp is enclosed in a vertical housing (see, for example, fig. 1), through which cooling air is blown. The spherical part of the bulb narrows the air-gap, and so increases the velocity of the air, precisely opposite the source of heat (the incandescent element).

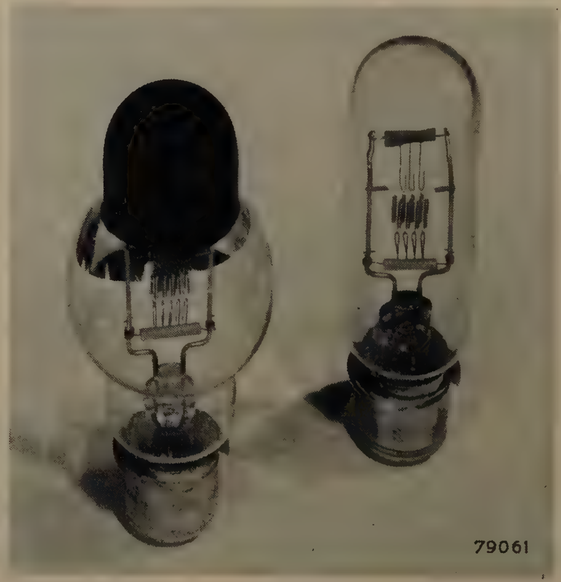


Fig. 4. The lamp employed in the 16 mm projector EL 5000 (left) and a conventional lamp for a 16 mm projector (right). The area covered by the filament of the left-hand lamp is rectangular, (same shape as the film gate), and the bulb of this lamp has a spherical form, part of which is internally silvered. Most of the material vaporised from the filament is deposited in the cylindrical top of this bulb, which is blackened to prevent light from escaping upwards.

---

[4]) For particulars of the design and properties of incandescent lamps for film projection, see Th. J. J. A. Manders, Philips tech. Rev. 8, 72-81, 1946.

One feature of the incandescent lamp employed in this projector is that the filament is rectangular (2 × 5 spirals), instead of square (2 × 4 spirals) as in the lamps usually fitted in projectors (see fig 4). To explain why the new shape is preferred, it is assumed for the moment that the condenser forms an image of the filament in the film gate (fig. 3*b*). (In reality, the image is formed at a point beyond the film gate, so as to prevent the structure of the filament from appearing on the screen.) Some of the light emitted by a square filament invariably falls wide of the film gate; therefore by making the contour of the filament indentical with that of the film gate, without increasing the total filament area, the amount of magnification involved in the formation of the image in the gate is reduced, so that angle u is increased. This enables the marginal rays to fill the objective more completely.

The objective

Objectives of relative aperture (diameter of lens divided by focal length) 1:1.6 are employed in almost all modern 16 mm projectors. The relative aperture of the objective in the EL 5000 projector however, is 1:1.3. The relative aperture $1/m$ and the angular aperture α already referred to are related by the equation [5]):

$$1/m = 2 \sin \tfrac{1}{2}\alpha$$

According to equation (1), then, increasing $1/m$ from 1:1.6 to 1:1.3 is equivalent, under otherwise identical conditions, to increasing the luminous flux by a factor of $(1.6/1.3)^2 \approx 1.5$.

As in most modern equipment of this kind, all the refracting surfaces of the objective are coated (bloomed) with a substance which reduces reflection [6]).

Shutter losses

One of the principal sources of light loss is the shutter. This is a rotating disc with two sectors, namely the frame sector and the flicker sector. The frame sector intercepts the light during the period in which the film is moved forward to bring the next frame, or picture, into the film gate. Unfortunately, the remainder of the total period available per frame cannot be employed exclusively for projection, since during it the beam must be cut off once more to prevent flicker (doubling the frequency). This second interception is effected by the flicker sector [7]). The duration of the two interceptions must be the same, or, in other words, the frame and flicker sectors must be exactly equal in width, since the slightest disparity between them will give rise to a perceptible and annoying flicker in the picture. This is due to the fact that if the two sectors are not properly matched, the fundamental frequency of exposure will still equal the frame frequency, instead of being twice the latter, as is the case when the sectors are identical.

Accordingly, if the frame-shift period is equal to p % of the total period per frame, a light loss of $2p$ % occurs; hence the importance of keeping the moving period as short as possible.

The loss occurring in the shutter is all the greater owing to the so-called "covering" angle (*fig. 5*). In theory, the beam should be completely cut off at the exact moment when the film starts to move, and should not be re-exposed until the film is stationary. Accordingly, the frame sector and also the flicker sector should be an angle β (covering angle) wider than the angle γ corresponding to the frame-shift period. Fortunately, however, practical experience has shown that the true situation is more favourable than the above argument suggests. Since the speed of the film is low at the start and end of each moving period. it is possible to "steal" part of the theoretical shutter width. To do so we determine by experiment exactly how narrow the shutter sectors can be made without permitting any part of the film movement to show on the screen. If the particular intermittent device employed is a Maltese cross mechanism with a four-slot cross (see below), as it almost invariably is in the case of 35 mm film, the moving period is governed entirely by this mechanism and gives $p = 25$ %. It is then possible to "steal" almost as much as is lost owing to the covering angle; hence the total shutter loss is about 50 %.

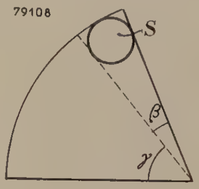


Fig. 5. The covering angle associated with the shutter sectors. The film must not start to move until the light beam (section S) is completely cut off. Hence the sector must be an angle β wider than the angle γ corresponding to the frame-shift period.

As we shall see later on, one attractive feature of the intermittent mechanism employed in the present projector is that it gives a free choice of the frame shift period. This is limited only by the amount of acceleration which the film and the mechanism will stand. We adopted $p = 16\frac{1}{2}$ %, which corresponds to a theoretical value of the frame angle (γ) of 60°. However, the covering angle β in our projector is so small as to enable us to "steal" more from the width of the shutter sectors than is added by this angle. Hence each sector is required to span only 50°, and the shutter loss is thus limited to 28%.

---

[5]) It should be borne in mind that the well-known paraxial formulae are not applicable, owing to the fact that the angles made by the light rays with the optical axis are not small.

[6]) The objective is manufactured by the N.V. Optische Industrie "De Oude Delft", of Delft, Holland.

[7]) See J. Haantjes and F. W. de Vrijer, Flicker in television pictures, Philips tech. Rev. **13**, 55-60, 1951/52.

The covering angle is minimized firstly by positioning the shutter where the beam diameter is smallest that is, immediately behind the film gate, and secondly by making the distance between the optical axis and the shutter axis relatively long. The shutter is therefore large, and has a high moment of inertia; hence it also performs the function of a flywheel which is required to ensure that the central driving spindle rotates smoothly. Moreover, blades fitted to this large shutter convert it into a fan (*fig. 6*), which despite a relatively low speed of rotation (1440 r.p.m.), provides sufficient cooling air for the projector lamp, the film mask (border of the film gate) and the film itself. This method of cooling is superior to the small fan driven at 5000-6000 r.p.m by a separate motor, which is used in many projectors, in that, apart from the eliminiation of the extra motor, it is virtually noiseless (no tendency to "whine").

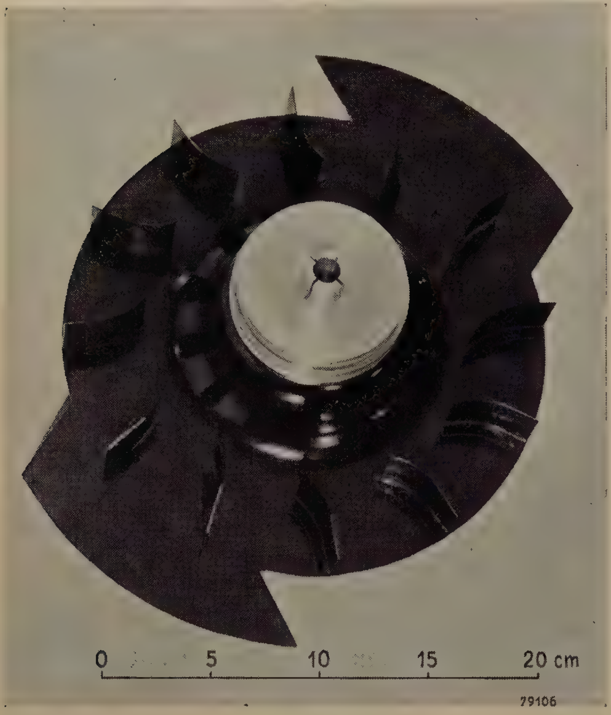


Fig. 6. The two-sector shutter. The frame sector and the flicker sector each cover 50°. The sectors are so shaped that the interception of the beam takes place parallel to the (long) side of the film gate, and therefore at a slightly higher speed than would be possible if sectors with truly radial edges were employed. Note also the large overall diameter, which enables the shutter to be used as a flywheel, and also the vanes which enable it to be used as a fan.

The intermittent mechanism

Principle

The intermittent mechanism, which enables us to procure the very short moving period already referred to, differs in many respects from the types usually employed in film projectors (a Maltese cross mechanism or a cam-driven claw mechanism). The principles of these conventional mechanisms are explained in *fig. 7*. The principle adopted for the projector that we are now considering may be understood from the drawing and photographs of *fig. 8*.

The intermittent sprocket, that is, the sprocket which feeds the film forward, is mounted on one end of a spindle. At the other end of this spindle is a disc, which has projecting pins spaced evenly all round its periphery. These pins fit into the grooves of a cam mounted on the main spindle, and rotating with the spindle at a constant speed, which, in terms of revolutions per second, is equivalent to the number of frames per second. It will be seen from the shape of the grooves, shown in fig. 8, that the frame-shift period corresponds to the angle δ through which the grooves move in advancing the pins by one position. With each individual pin-movement, the intermittent sprocket turns sufficiently to feed the film forward exactly one frame-length.

The associated accelerations are governed by by the shape of the grooves over that part of the cam corresponding to the angle δ. A suitable choice of this shape has enabled us to procure an angle δ of only 60° without imposing undue strain either on the film or on the mechanism, and thus to limit the light loss in the shutter, as explained in the preceding section, to only 28%. The acceleration of the film is then such as to preclude any further narrowing of angle δ (that is, any further shortening of the frame shift period).

Another valuable feature of this mechanism is that it gives freedom of choice as to the number of teeth on the intermittent sprocket. If the number of pins protruding from the disc be n, the angle through which the intermittent sprocket rotates with each full revolution of the cam is $360\,°/n$. Taken at the periphery of the intermittent sprocket, this angle must correspond to the frame pitch (length of film per frame); hence the circumference of the sprocket should be n times the frame pitch. Similarly, since 16 mm film contains only one perforation per frame, the sprocket should have n teeth. In the case here considered, $n = 12$. Accordingly, we have on the sprocket 12 teeth, which engage with the film perforations about six at a time; hence the teeth, as well as the sprocket holes in the film are subject to only a small amount of wear. Moreover, experience has shown that films whose perforations are so damaged that a claw mechanism will not properly engage with them, can be shown on this projector without any trouble at all, by virtue of the mechanism described here.

Defects in spacing

Mechanisms involving grooved cams operating on the above principle are widely used in automatic machines (e.g. bottle-filling machines), and have been attempted in cinema projectors. The novelty of the present intermittent mechanism lies in its structural design, which turns entirely upon the fact that the mechanism must be exceptionally accurate in every detail if it is to produce a stable picture on the screen. Unlike the claw mechanism, which returns to its starting point after feeding forward only one frame, the grooved cam mechanism does so only after advancing the film a distance of twelve frames. Inaccuracies in the distribution of the pins on the disc and of the teeth on the intermittent sprocket and also eccentricity of this sprocket are to some extent unavoidable. Though slight, these defects tend to affect the register of the intermediate frames so these do not all appear in exactly the same position; hence a kind of regular quivering, or dancing, of the picture results, which is all the more noticeable precisely because of its regularity. When once this effect is noticed, it continues to engage the attention of the observer and is therefore most irritating. To avoid it, the above-mentioned inaccuracies must be be kept to a minimum. For example, experience has shown that the Maltese cross mechanism employed in conjunction with 35 mm film, produces the familiar dancing pictures at a frequency of 6 per second (24 frames per second and a four-pole Maltese cross) if the wobble of the intermittent sprocket exceeds 1/100 mm. For 16 mm film, which requires a greater magnification, the wobble tolerance is smaller by a factor of 2.5, viz. about 4 μ. This fine tolerance can be attained in the manufacture of the intermittent sprocket. The teeth of the sprocket can also be cut to a sufficiently accurate *spacing*. Conventional methods, however, proved unequal to the task of accurately spacing the *pins* on the disc S_2 (fig. 8) of the new intermittent mechanism.

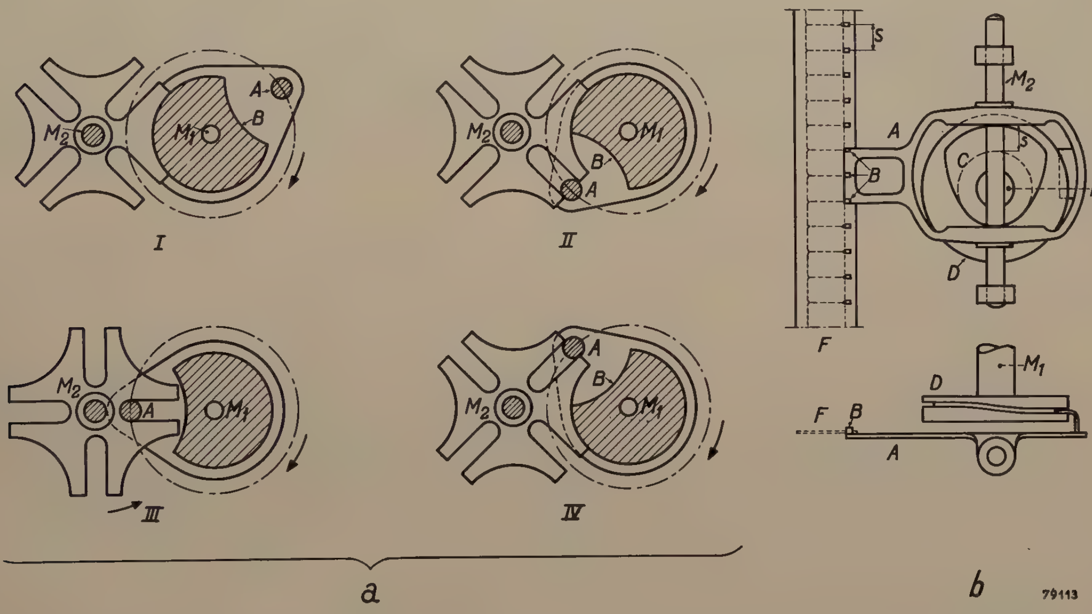


Fig. 7. Principle of the Maltese cross movement and the claw mechanism for intermittent film movement.
a) The Maltese cross movement in four consecutive positions. The end of spindle M_1, which rotates at a uniform rate, carries a cam with pin A and stop disc B. On spindle M_2 are the Maltese cross (here containing four slots), and the intermittent sprocket (not shown) which engages with the film perforations. In position *I* the Maltese cross is stationary; in position *II* the striker pin is just entering the cross, and the latter is starting to move; in position *III* the cross has been rotated through 45° and is now moving at its maximum velocity; in position *IV* the cross has completed a 90° rotation and is again stationary.
b) The claw mechanism. The arm (A) has a number of claws (B) (in this case three), which engage with the perforations of the film (F). A rotating cam (C) moves the arm up and down, the stroke of this movement being equal to the frame pitch. Cam C and the grooved cam (D) are mounted on the same spindle (M_1). The grooved cam turns the arm to and fro through a narrow angle about spindle M_2, so that the claws engage the film during the downward stroke, but miss it during the upward stroke.

A method was then adopted which enables correct pin-spacing to be effected automatically during the assembly of the disc, provided that each part is individually accurate. The pin-disc assembly seen in fig. 8 comprises a flat disc surrounding which 12 long and 12 short pins are arranged alternately, in intimate contact with each other and the outside of the disc. The pins are hollow, to reduce the mass of the assembly and are held in position by a clamping ring. The merit of this arrangement is that it enables very slight (unavoidable) variations in respect of the correct pin diameter to be compensated by deformation distributed uniformly amongst the pins. At the same time, the free, projecting ends of the long pins are not perceptibly affected by this deformation. Measurements showed that the pin-spacing is thus maintained accurate to within 1-2 μ.

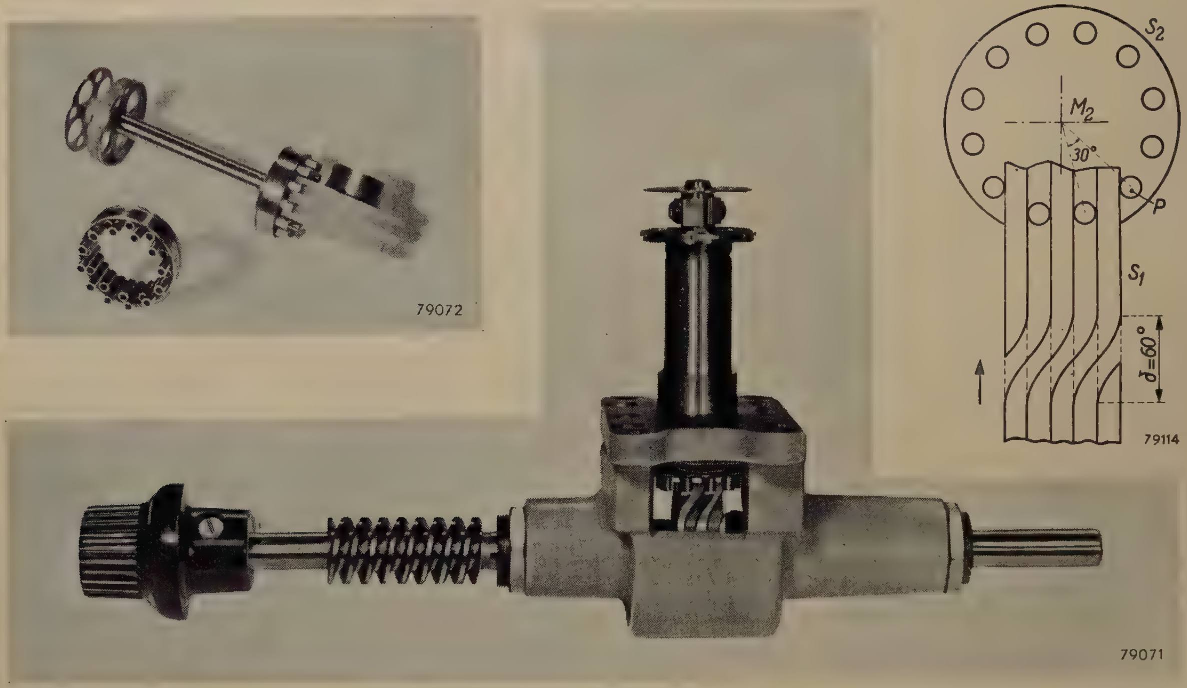


Fig. 8. The intermittent mechanism of the EL 5000. Top left: The most essential parts, viz. a spindle, with at one end the intermittent sprocket, and at the other a disc with projecting pins on its periphery. Meshing with these pins is the grooved actuating cam. The periphery of the disc with the pins positioned around it is included in the photograph to show how correct spacing of the projecting pins is achieved with the aid of shorter pins which do not project beyond the edge of the ring.
Top right: Diagram showing the action of the mechanism and the shape of the grooves of the cam (S_1). On the sprocket spindle (M_2) is the disc (S_2) with twelve projecting pins (P), which advance 1 position per revolution of S_1.
Below: Complete assembly. The pin disc, together with the grooved cam mounted on the main spindle, is completely enclosed in a housing (here cut open to show the interior) which is partly filled with oil. On the right-hand end of the main spindle will be mounted the shutter and a double belt pulley (fig. 6), which enables the mechanism to be driven at two different speeds. The worm on the main spindle drives the take-off sprocket and the take-up sprocket (see fig. 9 and fig. 12*a*). On the left is the inching knob used to move the film along by hand.

Another feature of this mechanism is that the grooved cam is made of nylon, which is very durable and resilient; so much so, in fact, that it is possible to make the width of the pin grooves about 10 μ *less* than the diameter of the pins and so preclude all possibility of play between the two. Continuous lubrication is ensured by housing the cam and the pin disc in an oil bath (fig. 8).

The shutter is mounted direct on the main spindle, which in this machine is at right-angles to the intermittent sprocket spindle. The main spindle carries a worm, the under-side of which engages direct with the bottom, or take-*up*, sprocket. The top part of this worm drives a large intermediate wheel, which actuates the top, or take-*off*, sprocket (*fig. 9*). In fact, the top and bottom sprockets both rotate at a constant rate, the one feeding, and the other ejecting the film (for a general view of the film lacing path, see fig. 12*a*). A double belt pulley enables the intermittent mechanism to be operated at two different speeds (18 frames per second for silent film and 24 frames per second for sound film).

Concluding this general description, we may

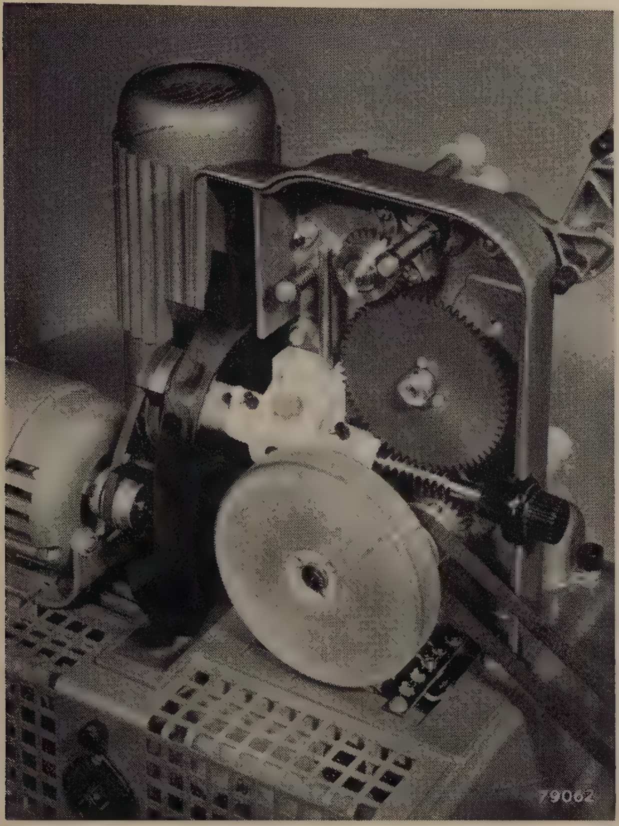


Fig. 9. Drive of the projector EL 5000. The main spindle, driven by the motor, passes through the housing of the intermittent mechanism, the top part of the worm on this spindle actuates a large intermediate wheel to drive the take-off sprocket, and the bottom part of this worm engages direct with the take-up sprocket (see fig. 12 *a*). The flywheel on the sound drum spindle can be seen in the foreground.

summarise as follows those features of the intermittent mechanism which are most relevant to its task: these are, a free choice of the frame shift period (which can hence be made very short) and of the exact way in which the film is accelerated, as well as of the number of teeth on the intermittent-sprocket; robustness and simplicity of design combined with accuracy, and, last but not least, efficient lubrication.

Comparison with the claw mechanism and the Maltese cross mechanism

To complete the description so far given it is necessary to add some remarks on the subject of the claw mechanism and the Maltese cross movement (fig. 7).

As regards the former a few words will be sufficient. Although in principle the claw mechanism is very accurate (see "Defects in spacing"), so much so, in fact, that almost all modern cinematograph cameras are equipped with it, none of the existing forms of this mechanism are suitable for use in a professional projector, for the following reasons. Firstly, these mechanisms invariably operate with very few claws, usually, in fact, only three, which between them must absorb all the wear arising from contact with the film. From this point of view alone, then, the life of our mechanism (12 teeth) will be four times as long. Secondly, the claws engage with the film perforations only three at a time, as compared with six at a time in our mechanism; hence the claws soon damage the perforations. Moreover, they fail to impart the correct movement to a film which has been so damaged. Thirdly, the claws must be light in order to perform the rapid reciprocating movement required of them, and this lightness is not compatible with robust construction. Finally, the lubrication of a claw mechanism presents a very difficult problem.

With regard to the Maltese cross mechanism, it can of course be adapted to 16 mm film, but the mechanism is then deprived of much of its attractive simplicity. To explain why the ordinary form of Maltese cross mechanism is never employed for 16 mm film, we shall now describe it more fully.

A diagram showing the principle of this movement (here supposed to be of the four-pole type) will be seen in fig. 7. One condition which every Maltese cross movement must fulfil in order to operate smoothly is that the striking pin shall be travelling in the direction of the centre line of the slots in the Maltese cross at the precise moment of entering or leaving one of these slots (which are stationary at such a moment). In the diagram shown in *fig. 10*, the striking pin is describing a circle about M_1; M_2 is the spindle of the Maltese cross, A_1 the position of the striking pin on entering, and A_2 the position of this pin on leaving a slot in the cross. It will be seen that angles α_1 and α_2 of the quadrilateral $M_1A_1M_2A_2$ are right angles; hence $\beta_1 + \beta_2 = 180°$. In the case of a cross with m slots, $\beta_2 = 360\,°/m$. The period in which the pin is engaged with the cross, i.e. the moving period, is proportional to β_1; to procure a short moving period, then, we must make β_1 small and consequently β_2 large. A three-slot Maltese cross furnishes the smallest angle β_1, viz. 60°, which then corresponds to a moving period as short as that of the new mechanism employed in the present projector. However, such a cross gives no freedom of choice as to the precise way in which the film is accelerated; in fact this is so violent that the forces involved soon cause damage to the film and to the mechanism. Accordingly, a Maltese cross should contain at least four slots, the moving period then being 25 % of the total period per frame, and the light loss in the shutter about 50 % of the total flux. In the case of the four-slot cross, the circumference of the intermittent sprocket, when mounted direct on the spindle of the cross, must be equivalent to four frames of

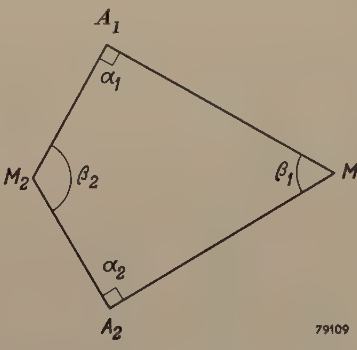


Fig. 10. Diagram showing method of determining the moving period of an *m*-pole Maltese cross mechanism. The striker pin rotates about M_1 at a constant rate. M_2 is the centre of rotation of the Maltese cross. Points A_1 and A_2 are the respective positions of the striker pin on entering and on leaving the cross; hence $\beta_2 = 360\,°/m$. Since the principle condition of smooth operation is that $\alpha_1 = \alpha_2 = 90°$, angle β_1 (and, therefore, the frame shift period) increases with *m*.

the film. For 16 mm film, then, the sprocket required would be so small that it would force the film into a very tight loop, which would tend to unstick any splices in the film. A still more serious disadvantage of such a sprocket is that it can accomodate only four teeth, which can engage with the film only two at a time; hence the film perforations would soon be damaged and the teeth soon worn.

Freedom of choice as to the number of teeth on the intermittent sprocket can be obtained by mounting the sprocket on a separate spindle geared-down to the spindle of the cross. However, apart from the extra spindle, this involves an extra gear, which, in view of the spacing defects already referred to, must be very accurately machined. The moving period can be shortened by varying the speed of rotation of the striking pin in pulses, so that it rotates quickly when actuating the cross, but slowly when clear of it. On the other hand, the practical application of this method involves one (or two) extra spindles. It will be seen, then, that a Maltese cross mechanism of approximately the same efficiency as our grooved disc mechanism would be very complex.

Details of the film threading

Automatic tensioning of the film

The bottom, or take-up, sprocket feeds the film to the take-up spool: this is driven, as usual, by a drive which slips when necessary to adapt the speed of rotation to the varying diameter of the film reel. It is desirable that the tension of the film between the sprocket and the take-up spool be maintained virtually constant in the region of 160 grams, since at this tension the film is wound just taut enough. The film tension, being equal to the frictional driving torque divided by the radius of the film reel, will be constant if one of these factors is made proportional to the other. To fulfil this condition a very simple attachment is employed. This is an arm secured to the projector frame by a hinge and carrying at its free end the spindle for the take-up spool (*fig. 11*). The free end of this arm is supported solely by an endless belt of braided cotton, which passes round a pulley at one end of the spool spindle and constitutes a slipping drive for the take-up spool. The friction between belt and pulley, and hence the driving torque, increases with the weight of the film reel. Now, by adopting a suitable pulley diameter and positioning the driving belt and the spool arm at suitable angles to the vertical, it is possible to make the driving torque proportional to the radius of the film reel, and their ratio equal to the desired film tension. In the projectors now in production, the film tension changes only from an initial 170 grams for an empty spool to a final 150 grams for a full spool (600 m of film). The variable friction clutch to be found in most projectors has been abolished in this system, and with it has gone the possibility of incorrect adjustment. This is an even greater advantage from the point of view of sub-standard film than from that of ordinary 35 mm stock, the former being by far the more fragile of the two owing to its relatively small size and less durable (though non-inflammable) quality Moreover, it is probable that projectors for sub-standard film will sometimes be operated by untrained, or inexperienced persons.

In the case of the feed spool, the problem of film tension was solved in a similar manner. The torque set up by frictional forces in the bearing of the spindle carrying this spool decreases with the weight, that is the radius, of the film reel, and so maintains the film tension virtually constant. A frictional torque consistent with the desired film tension is procured by employing a bearing of the appropriate diameter.

Threading the film

In view of the possibility already referred to, that the projector will be operated by persons not specially trained for the purpose, the process of threading the film has been made as simple as possible. The mere turning of a handle clears the film path of all obstructions (see *fig. 12*), and positions the pad rollers so that the film can be threaded in taut. When once the film is so threaded, the handle should be turned back to its original position, and the projector is then ready to operate. Loops of the exact length required to permit of the necessary intermittent movement through the film gate are formed automatically above the gate and beyond the intermittent sprocket. The length of the bottom loop (beyond the intermittent sprocket) is especially

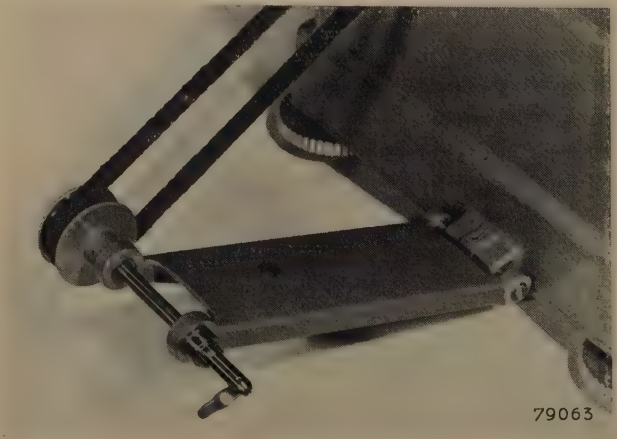


Fig. 11. Drive and mounting of the take-up spool. One end of the spool arm is hinged to the projector, and the other end, carrying the spool spindle, is supported solely by the belt of braided cotton which drives the spindle. This belt slips whenever necessary to adapt the speed of rotation of the spool to the diameter of the film reel: the film tension remains almost constant at approximately 160 grams.

critical, since it governs the distance between the particular frame exposed in the film gate at a given moment and the part of the sound track that will be scanned at the same moment. In the case of 16 mm film, the distance between the two should always be 26 frames, this being the standard space between a frame and the associated sound. Proper length adjustment of the bottom loop is all the more important in the case of 16 mm film because here the number of frames per unit length is a factor of 2.5 more than in 35 mm film. Hence an error of one or two cm in the length of this loop causes far more asynchronism in a 16 mm, than in a 35 mm film.

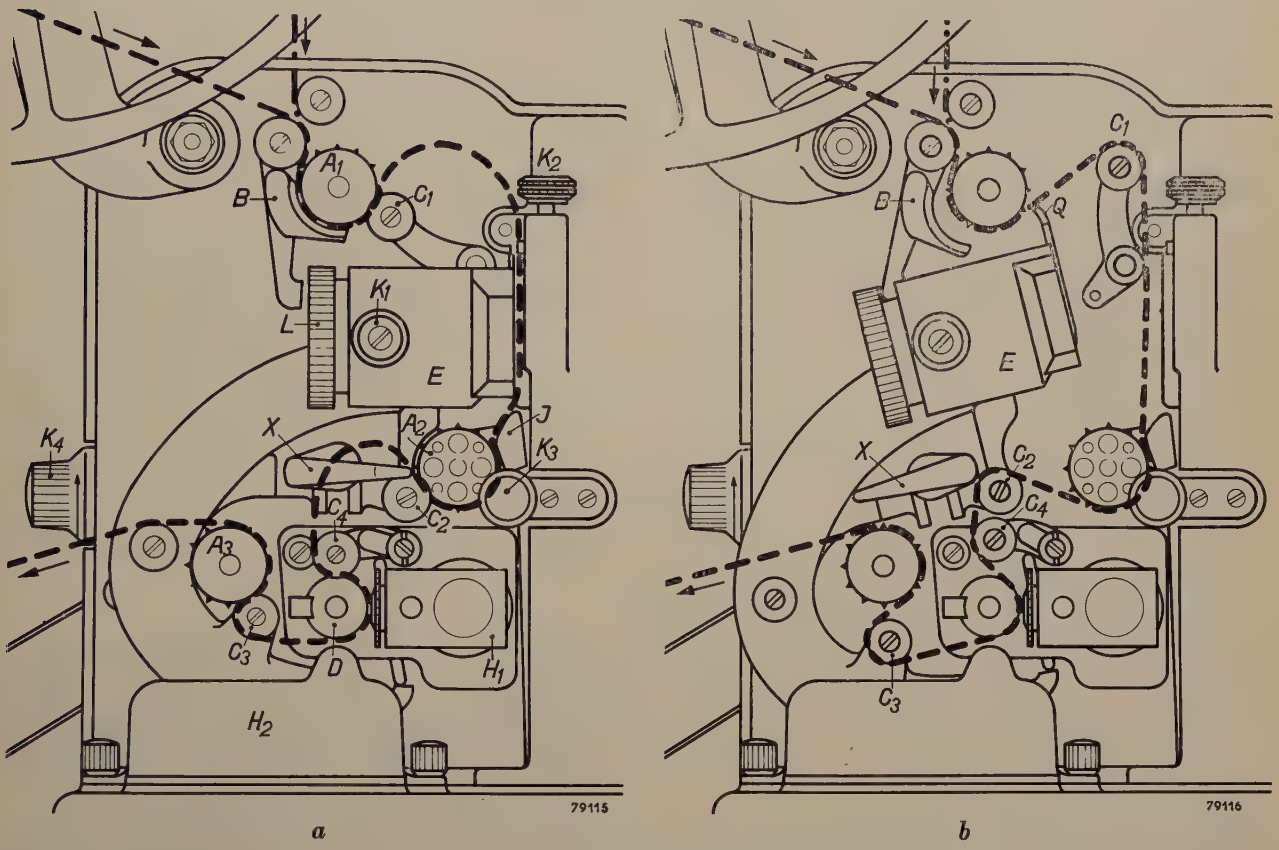


Fig. 12. *a*) General view of the film lacing path. The take-off sprocket A_1 draws the film from the feed spool (at this point the film may follow either the broken line or the chain-dotted line, depending on how it has been wound). Guide shoe B and pad roller C_1 prevent the teeth from slipping out of the film perforations. The intermittent sprocket A_2 draws the film intermittently through the film gate, between the fixed guide and the spring guide. (The fixed guide Q will be seen in (*b*)). Next, the film travels over the spring-loaded pressure roller C_4 and round the sound drum D, and thence to the take-up sprocket A_3 with pad roller C_3, which feed it to the take-up spool. E is the objective mount, L the objective itself and K_1 the objective locking screw. K_2 is the knob for varying the spring pressure on the film in the gate, and J the framing shoe with adjusting knob K_3. H_1 is a cover to protect the exciter lamp and optical system of the sound-head, and H_2 similarly protects the photo-electric cell. K_4 is the inching knob and X the handle to clear the projector for film threading.
b) The projector cleared for film threading, with film threaded in. The handle X has been turned, so that the objective holder E, together with the fixed guide Q attached to it are tilted back. At the same time, the shoe B near the take-off sprocket and pressure roller C_4 on the sound drum, are lifted and so positioned that the film can be threaded *taut* along the proper path. When handle X is turned back to its original position, loops of the correct lengths will be formed above the film guides and beyond the intermittent sprocket (see (*a*)). Sound and picture are thus synchronized automatically.

The film gate

The film is held stationary in the film gate between two straight guides (the runner plates), one in front and one behind. One of them, which is spring-loaded, may be described as the spring guide, and the other as the fixed guide (fig. 12 *b*). Facilities for adjustment of the spring pressure by the operator are provided for the following reason. Films, especially new ones, deposit dirt on the film guide. This dirt usually accumulates rather rapidly and may so increase the spring pressure that the latter causes the film to break. By means of the adjustment provided, the operator can reduce the pressure whilst the film is running, and so need not interrupt a programme in order to remove the dirt.

The fixed guide, unlike those in ordinary 35 mm projectors, is attached to the objective holder: this ensures that the film will remain in focus even if handle X is not turned quite as far as it should be after the insertion of the film, and the objective is therefore slightly out of alignment (fig. 12*b*).

The framing adjustment

The system must be flexible enough to enable individual frames to be accurately registered in the film gate as and when required (framing). Framing is necessary to compensate for film shrinkage (new films are invariably longer than old ones) and for any frames positioned either too high or too low in relation to the sprocket holes, as may happen, for example, in the process of copying an original film. Framing adjustment can be obtained by making the length of the film path between the gate and the intermittent sprocket variable. Variability of the film path is procured by taking the film from film gate to intermittent sprocket over an adjustable shoe (J in fig. 12*a*); moving this shoe to the left lengthens the film path. One advantage of this system is that it keeps the picture centred on the screen during framing, unlike, for example, the widely used method of making the film gate itself, or more precisely the film mask mounted in this gate, adjustable. The last-mentioned method would here necessitate vertical re-adjustment of the projector after framing, which, for professional equipment, is inadmissible. We were able to adopt the more convenient solution by virtue of the fact that the framing travel required for 16 mm film is only about 1 mm. Far more range of adjustment is required for 35 mm film. In this case, the pitch of the sprocket holes is 1/4 of the frame pitch, therefore a badly positioned splice may cause the film to jump $\frac{1}{4}$, $\frac{1}{2}$, or even $\frac{3}{4}$ of a frame; hence the 35 mm framing system must be capable of correcting such variations. However, similar picture displacements do not occur in the case of 16 mm film, where the sprocket hole pitch is equal to the frame pitch.

The film path in the sound head

How to maintain the film speed constant within the sound scanning system has remained a problem ever since the advent of sound film. Vibrations are produced in the film owing to the discontinuous movement of the intermittent sprocket and also to the fact that the teeth of the take-up sprocket do not draw the film perfectly smoothly through the sound-head. These vibrations must be prevented from reaching the point at which the sound track is scanned. A conventional solution to this problem, viz. a rotating sound drum, is used. This is a drum driven by the film and coupled to a flywheel and therefore rotating at a very uniform rate. The sound track is scanned at a point on the drum [8]). The loops described by the film in passing round the different rollers act as a spring, and this spring, in conjunction with the combined mass of the flywheel and the sound drum, forms a mechanical filter through which the vibrations cannot pass. The lighter the spring and the greater the moment of inertia of the flywheel, the more effective is this filter.

A light spring is obtained when the loops in the film are loose, that is, when the tension of the film is low. Now, low film tension between the sound drum and the take-up sprocket can be procured by running the spindles of the drum and the associated pressure roller in ball-bearings lubricated with thin oil instead of grease, which permits of very free rotation.

The film is often threaded in an S-bend, to make it fit closely round the sound drum (*fig. 13a*).

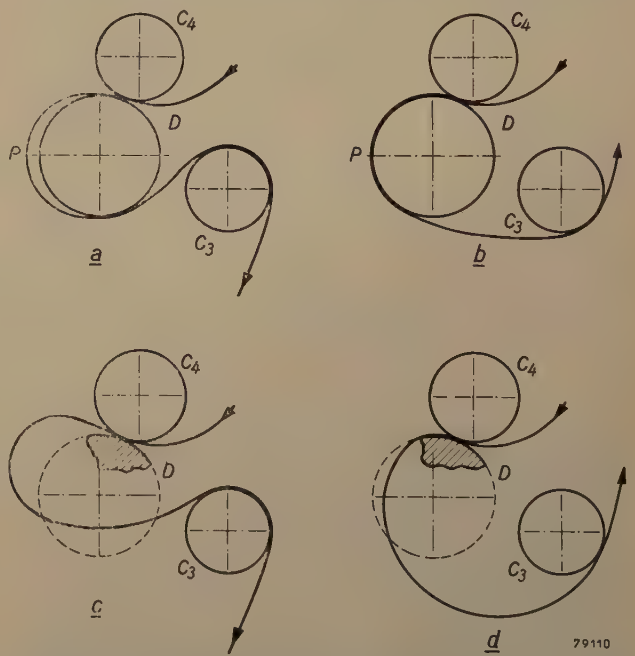


Fig. 13. The film lacing path in the sound-head. D is the sound drum, C_4 the pressure roller, C_3 the guide roller and P the sound scanning point.
a) Threading the film round the sound roller and the guide roller in an S-bend, in conjunction with low film tension, tends to cause the film to stand away from the sound drum precisely at the scanning point P ("breathing").
b) With the film threaded in a U-bend, it will remain in close contact with the sound drum at point P despite low film tension.
c) The loop which the stationary film would describe in case (*a*), if the whole of the sound drum other than the portion hatched in the diagram were removed.
d) The same, in respect of case (*b*).

---

[8]) See J. J. C Hardenberg, The transport of sound film in apparatus for recording and reproduction, Philips tech. Rev. **5**, 74-81, 1940.

Logical though this may seem, however, it is found that in fact, owing to the above-mentioned low film tension, the film then becomes detached from the drum precisely at the scanning point, and goes into almost imperceptible radial oscillation ("breathing"); its consequent failure to remain continuously in proper focus relative to the optical system gives rise to perceptible sound distortion. On the other hand, the film remains in close contact with the drum if taken round the other side of the guide roller to form a *U*-bend instead of an *S*-bend (fig. 13*b*). This apparently strange behaviour of the film is nevertheless to some extent understandable bearing in mind what shape the film would assume if the the whole of the sound drum other than the small portion actually in contact with the pressure roller were removed (fig. 13*c* and *d*).

The sound installation

Details of the scanning system

The light source employed for sound scanning is an illuminated slit, a sharp image of which can be focussed on the sound track of the film by turning the large, knurled ring shown in *fig. 14*. Moreover, the slit can also be adjusted so that it is at right-angles to the sound track (azimuth adjustment). The two adjustments are quite independent of each other.

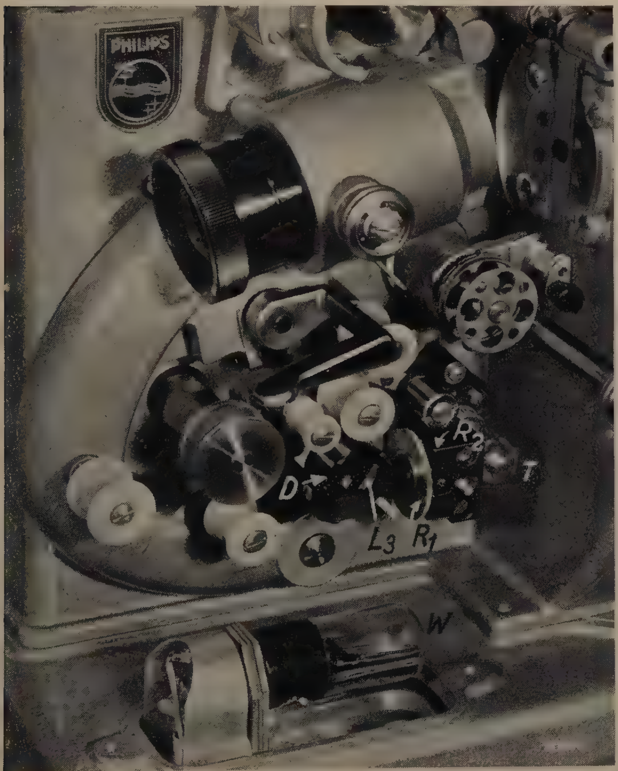


Fig. 14. The sound-head of projector EL 5000. *T* is the exciter lamp. Knurled ring R_1 is used to focus the slit-image on the sound track, and R_2 to adjust this image at right-angles to the track (azimuth adjustment). The plano-convex lens L_3, which projects the light down into the photo-electric cell *W*, is just visible in the photograph, inside the sound drum *D*. This photograph shows quite clearly that the guide rollers are not made of steel; they are made of nylon.

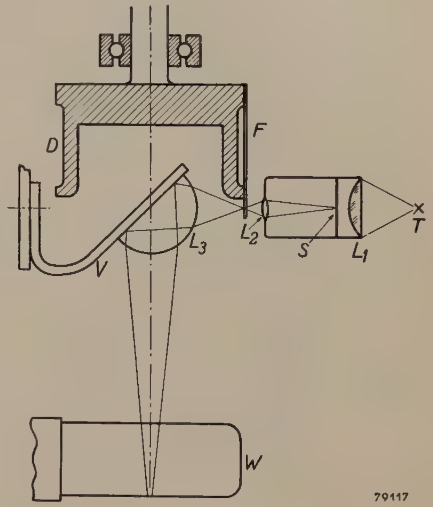


Fig. 15. Diagram showing the path of light-rays during scanning. The sound track of the film *F* projects beyond the edge of the sound drum *D*. *T* is the exciter lamp, *S* the slit, L_1 the condenser lens, L_2 a cylindrical lens and L_3 the plano-convex lens, silvered on the plane side; *W* is the photo-electric cell and *V* a support for L_3. In reality, the sound drum is mounted with its axis at right angles to the position shown, i.e. perpendicular to the plane of the diagram.

The lamp is exited by an oscillator supplying H.F. current (90 kc/s), and therefore produces no audible note. Since the output of the oscillator must in any case be extracted through an output transformer, there was, with the available power of 10 watts, a free choice as regards the current and voltage for the exciter lamp. In the case here considered, a low voltage and a heavy current were adopted, viz. 2.5 volts and 3 amps., which differ considerably from those employed in most sound-heads. Accordingly, the filament is a short, tough spiral of thick wire, producing no microphony.

Another point worth mentioning is that the exciter lamp may be switched over to a 50 c/s supply if a fault develops in the oscillator. The resultant hum is relatively slight, owing to the high thermal capacity of the thick filament.

The light emitted from the slit is projected downwards by a plano-convex lens, silvered on the plane side, into a photoelectric cell (*fig. 15*) secured to the mounting plate of the amplifier. The removal of two srews, one on either side of the apparatus, enables the whole projector to be lifted from the amplifier, leaving the photo-electric cell still in

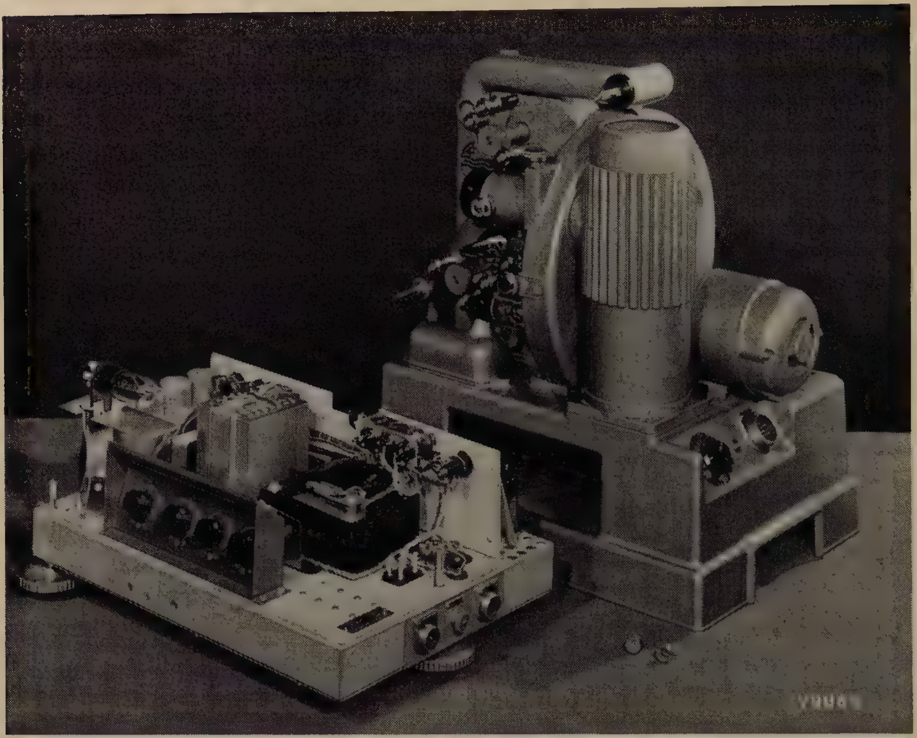


Fig. 16. The projector is completely detachable from the amplifier. Note the photoelectric cell at the top left-hand corner of the amplifier.

position (*fig. 16*). In this way, neither a photoelectric cell cable, or the associated plugs, which are apt to cause interference (crackle), are required.

The projector is so designed that, from the rear, the film is at the left-hand side and the sound track accordingly on the outer edge of the film; hence the sound scanning beam does not fall foul of the spindle of the sound drum (see fig. 15) and the plano-convex lens can be mounted inside this drum. Moreover, the lens is easily removed, for cleaning or inspecting the optical system.

The sprocket holes being along the inner edge of the film the latter can be inserted without being passed full-width across the sprocket teeth and so exposed to possible damage by these teeth.

Features of the amplifier

The sound amplifier gives an total output of 15 watts, with 3% distortion; this output is divided between two loudspeakers, together powerful enough for a hall seating 750 people. The high and low note response can be varied separately. A socket for a microphone (or gramophone) is provided in the amplifier, so that a commentary or explanatory talk can be given to accompany the pictures on the screen (see fig. 1). The volume of the sound proceeding from the film itself and of the spoken, or "live" commentary can be varied independently.

The design of the amplifier, like that of the remainder of the equipment described here, is very largely the outcome of a desire to make the projector as reliable as possible, in view of the fact that it is intended for professional use. In order to achieve this, the number of valves in the amplifier is quite considerable (i.e. 12, including the high-frequency supply unit for the sound track exciter lamp). This has enabled us so to arrange matters that in the event of failure of one or more of the components during the programme, the amplifier can still function, though with some deterioration in quality and (or) volume of the sound. In fact, the amplifier will operate with only six valves.

Only four different types of valve are employed; hence four valves in all constitute a complete reserve set. Each of them is a standard type of radio valve and therefore readily obtainable.

Further particulars

The projector is driven by a split phase motor, amply dimensioned and therefore unresponsive to variations either in voltage or in load. In this type of motor there is no commutator to cause sound distortion owing to sparking at the brushes. The motor is mounted on rubber to prevent the transmission of vibrations to the projector.

Suspended in the lamp housing is a vane, which cuts off the light passing to the film when the latter is at rest. When the projector is running, that is, when the film is moving, however, the draught of cooling air set up by the shutter-fan lifts the vane, thus admitting the light to the film. This vane is provided, not so much against the actual risk of fire, but to prevent overheating and damage to the film, should the lamp remain switched on with the film stationary.

A separate re-winder is provided (see fig. 1), which dispenses with the inconvenience of taking the projector out of service in order to re-wind a film.

Practical experience has shown that the projector described here fully satisfies the requirements imposed on it during design. The quality of the sound and picture depends, of course, also on the quality of the available films and regrettably few sub-standard films at present on the market are of really high quality as regards both picture and sound. However, the present trend of development is towards a considerable improvement in this respect.

Summary. For some time past, the Philips factories at Eindhoven have been producing a 16 mm sub-standard film projector (type EL 5000) designed for professional use. Two important features of this projector are its high light output (500 lumens) and its robust construction. A high light yield, that is, sufficient for the projection of 4 × 3 metre pictures, is obtained firstly by employing an objective of large relative aperture (1:1.3), secondly by carefully matching the lighting system with this objective, and thirdly by minimizing shutter losses. The shutter losses are reduced by means of an unconventional intermittent mechanism, which ensures a very short frame shift period. Another advantage of this mechanism is that an intermittent sprocket having a large number of teeth can be employed. This mechanism is described in detail. Shutter losses are further reduced (to 28%) by means of a special, large-diameter shutter. The size of this shutter also enables it to be used as a fan and simultaneously as a flywheel. Other details of the design discussed in this article are: automatic control of the film tension at the feed and take-up spools; automatic looping of the film, and automatic synchronization of picture and sound during the threading of the film; framing adjustment without de-centring of the picture on the screen; very low film tension before and after the sound-head, whereby a constant film speed is maintained at the scanning point; and, lastly, the sound scanning and amplification system itself.

VISUAL ACUITY IN CONNECTION WITH TELEVISION

by G. J. FORTUIN.

621.843.6

In 1951, Philips' Medical Department tested 228 persons of ages varying from 7 to 64 years in order to collect data concerning visual acuity and its relation to contrast and brightness. In this article the results of this research are discussed in connection with a study of the observation of television images.

The information obtained when looking at a two-dimensional, flat object, e.g. a photograph or a television picture, is determined on the one hand by the sharpness and contrast of the image itself, and, on the other, by the properties of the observer's eye and the conditions of observation (brightness, distance between the eye and the object).

The property of the eye that is decisive in this respect, viz. the visual acuity, can be expressed quantitatively by the angle D_0 subtended at the eye by the smallest perceptible detail. This angle is frequently assumed to be approximately one minute of a degree, but it is highly misleading to attribute a fixed value to D_0. As will be seen, D_0 depends not only on the luminance, but also to a very considerable extent on the contrast. Furthermore, there may be individual differences, which may be accentuated by insufficient correction of refraction errors (wrong spectacles, or no spectacles at all). Moreover the *age* of the observer greatly affects the value of D_0.

In television the smallest detail in the vertical direction is determined by the scanning width: in the case of a 625-line, 30 × 40 cm picture this width is approximately 0.5 mm. The smallest detail in the horizontal direction is determined by the highest modulation frequency, and it can be assumed that it is of the same order of magnitude. If d be the size of the smallest detail, then the maximum distance at which it can be discerned, is $a = d/D_0$. With $d \approx 0.5$ mm and $D_0 = 1'$, it follows that $a = 1.8$ m. This, of course, is only an approximation, for a is actually a dependent on all the above-mentioned factors which affect the value of D_0.

It is evident that as a rule the observer wants to watch the screen from a distance at which it is just possible for him to discern the smallest details with smallest contrasts. As the contrast between the lines of the image and the spaces in between is greater than the smallest contrast the observer wants to see, the scanning lines themselves will also be clearly distinguished. Experience has shown that viewers do not object to this.

It is important in this connection to consider more closely the factors influencing the value of D_0.

In 1951 Philips' Medical Department carried out extensive research into the relationship between visual acuity, contrast and brightness in connection with the age of the observer [1]), and the results obtained in this work are applicable to the problem under discussion. The 228 persons tested (of ages varying from 7 to 64 years) comprised school children, applicants for jobs, and in the higher age categories, employees of the Philips works. To permit conclusions of practical value to be drawn from the tests, all these persons were subjected to the examination in their normal, every-day circumstances, i.e. no attempts were made to improve their eyesight, for example, by means of spectacles. The original object of the investigation was to examine the influence of age on vision during very delicate manufacturing operations (assembling of radio tubes) and to what extent this can be compensated by raising the level of illumination .The tests partly confirmed earlier results about the relationship between visual acuity, contrast and brightness [2]), which could now be combined into a single empirical formula, also containing a factor concerning the influence of the observers' age on his visual acuity.

Testing procedure

Since an object can be fairly easily recognized according to its shape (this is illustrated, e.g. by the varied shapes of printed letters), it is desirable in an investigation into visual acuity to use some kind of standardized object. In the present case a so-called Landolt ring was used for this purpose (*fig. 1*).

The rings are on a white panel (*fig. 2*) which is divided into 16 × 11 squares, in each of which a black or grey paper Landolt ring is pasted. There are eight different ring positions (opening at top, bottom, left, right, or at 45° in between). In each vertical column the outside diameter of the rings diminishes from top to bottom according to a geometrical progression from 50 to 2.2. In each horizontal row the contrast diminishes from right to left according to an arithmetical progression from 0.94 to 0. The contrast is determined by the difference in luminance $L_a - L_r$ between background and ring divided by the background luminance L_a; alternatively if background and ring receive equal amounts of light, the contrast is determined by the difference in reflective power $\varrho_a - \varrho_r$ between background and ring divided by the reflective power ϱ_a of the background:

---

[1]) G. J. Fortuin, Visual power and visibility, Diss. Groningen, 1951; see also Philips Res. Rep. **6**, 251-287, 347-371, 1951.

[2]) See e.g. P. J. Bouma, Visual acuity and speed of vision in road lighting, Philips tech. Rev. **1**, 215-219, 1936.

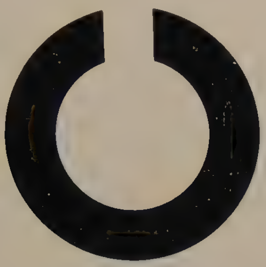


Fig. 1. Landolt ring. The aperture and the width of the ring are equal to one fifth of the outside diameter.

$$C = \frac{L_a - L_r}{L_a} = \frac{\varrho_a - \varrho_r}{\varrho_a}.$$

The visual acuity in this Landolt ring test is determined by the reciprocal value of the angle (expressed in minutes of arc) subtended by the aperture of the ring, in the case of the smallest ring whose position the observer is able to ascertain. If this angle be D_0, then $1/D_0$ represents the visual acuity. The test supervisor indicates, by means of an arrow-head (see fig. 2), one ring at a time. *Fig. 3* shows how the person under test responds by placing a rotatable ring into a position corresponding with that of the ring indicated on the panel. A correct answer is signalled electrically (by a white disc appearing in front of an aperture, visible only to the investigator).

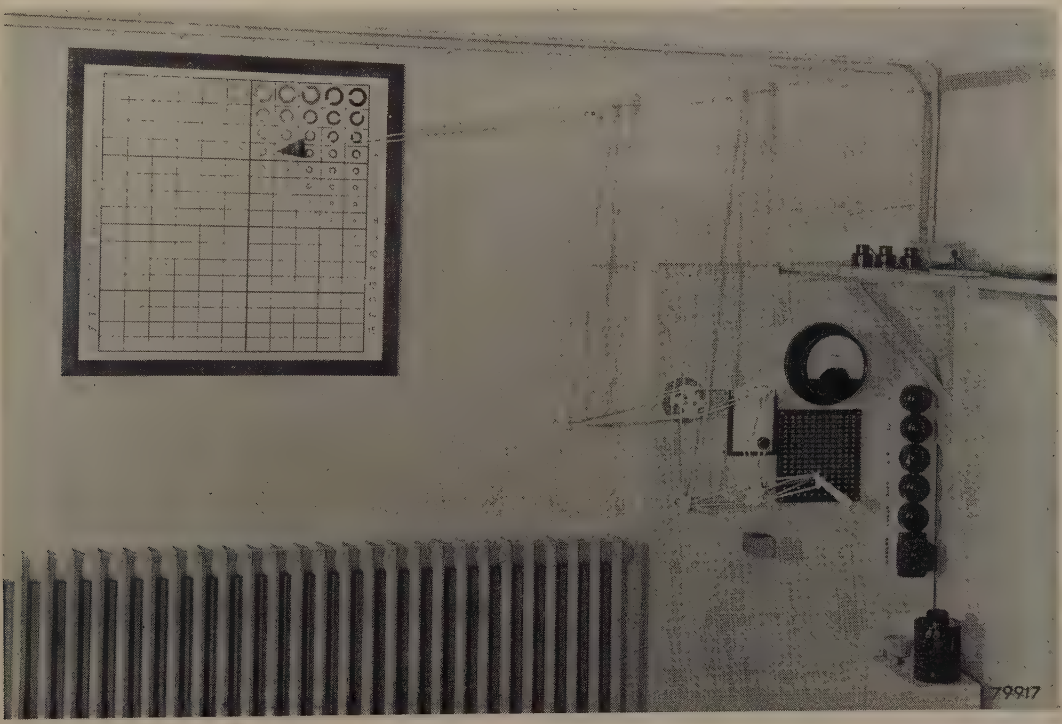


Fig. 2. Equipment for examining visual power. The panel on the left is divided into 11 × 16 squares, each of which contains a Landolt ring of known size and contrast value, but set in an arbitrary orientation. Each time the arrow points at one of the rings; the person under test, who stands at a distance of 5 m, has to try to distinguish its orientation. On the right, the semi-automatic recording device.

Fig. 2 also shows how this signal is produced. The movement of the arrow-head is reproduced on a smaller scale by means of a parallelogram linkage to a pin moving across an insulating board with holes and to a needle moving across a sheet of paper in a holder. Both the insulating board and the sheet of paper are small-scale reproductions of the ring panel. Each hole contains a metal bush which is connected to one of the eight contacts under the rotatable ring manipulated by the person under test. As soon as the pin is pressed into the hole, a circuit is completed provided the position of the rotatable ring corresponds with that of the paper ring indicated by the arrow-head.

When the answer is correct, the examiner presses the needle in, thus punching a hole in the paper and recording the result (*fig. 4*).

The final result is then ten figures [3] for each test person at various luminance values L_0 (four in our case) of the field of vision, viz. the numbers of the last ring in the various columns for which a correct answer was given. The value of D_0 was taken as the geometric mean of the angles relating to this last ring and the next smaller one. For each value of D_0 there is an accompanying value of the contrast C_0. Together L_0, D_0 and C_0 are the factors determining what is termed the "threshold value". In the investigation the various threshold values per test person were studied and compared individually.

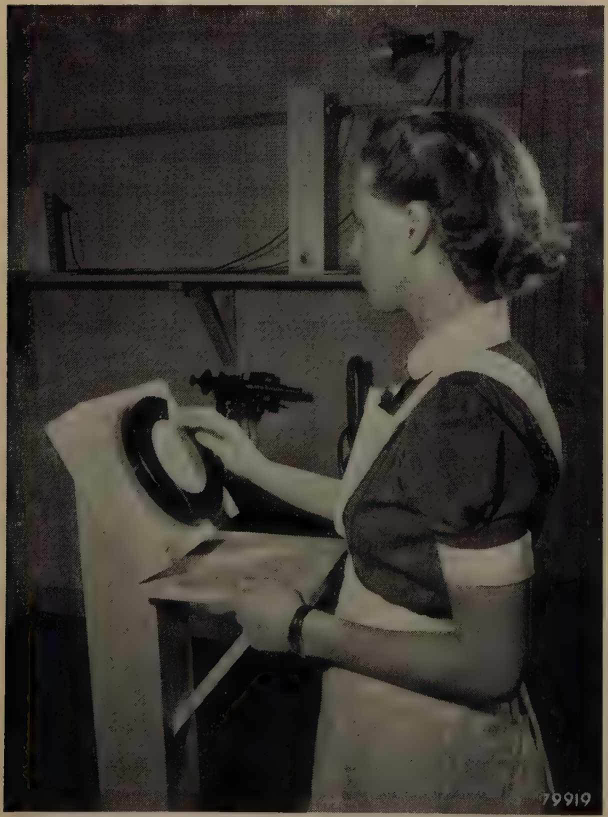


Fig. 3. The person under test answers the question by placing the rotatable ring in the same orientation as the one indicated.

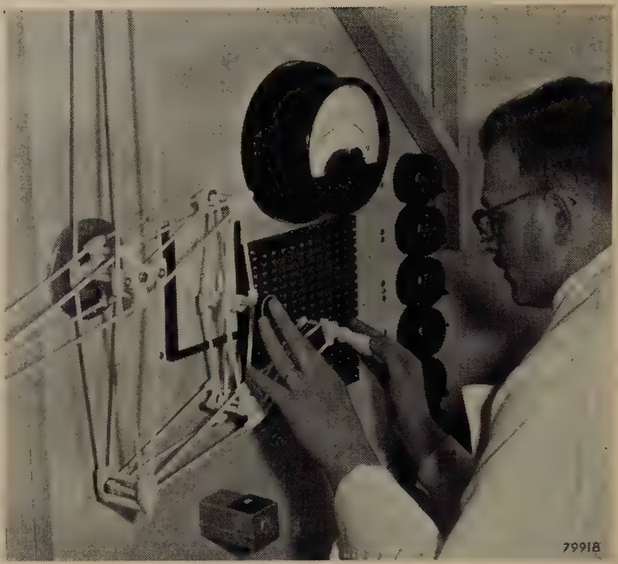


Fig. 4. The examiner at the instrument panel. Top: the voltmeter for checking the voltage of the lamps. With his left hand the examiner presses the needle down. Below: the electric signalling device which shows if the answer is correct.

Test results

The average result of all persons tested can be represented by a three-dimensional diagram, where D_0 is plotted as a function of C_0 and L_0 (*fig. 5*). It will be seen that D_0 diminishes (and consequently the visual acuity increases) at a given luminance as the contrast increases, and at a given contrast value as the luminance increases.

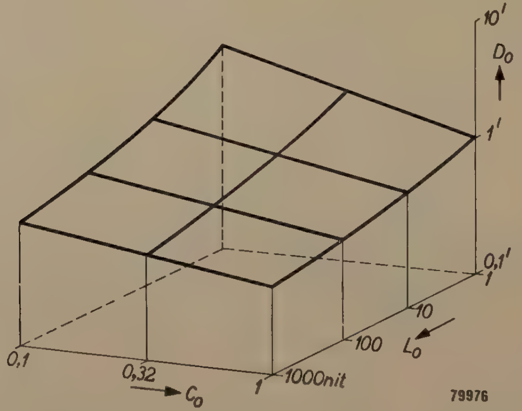


Fig. 5. The average result for all persons, represented in a three-dimensional diagram. It shows the relationship between the contrast value C_0 of the test object, the luminance L_0, and the smallest detail perceived U_0.

The surface in fig. 5 can be expressed by the empirical formula:

$$\log D_0 = 2.17 \frac{1.57 - \log C_0}{\log L_0 + 3.95} - 0.79 ;$$

where $\log D_0$ represents the average value of $\log D_0$

---

[3] In the extreme left-hand column the contrast was zero (paper of rings same colour as background). This column was not taken into consideration in the tests.

(for given C_0 and L_0) for all tested persons [4]). For each individual the same formula applies, viz.;

$$\log D_0 = 2.17 \frac{1.57 - \log C_0}{\log L_0 + 3.95} - \log G, \quad . \quad . \quad (1)$$

in which the value of G varies according to the individual tested.

G is a direct measure of the visual power. If in the case of two persons the value of G for the one is n times as large as that for the other, the former will, under equal conditions, be able to discern details that are n times as small [5]). Physically, $0.1\,G$ is the reciprocal value of the angle at which the observer is just able to see the opening in the Landolt ring at given combinations of C_0 and L_0, i.e. when the term of formula (1) containing C_0 and L_0 is equal to 1 (e.g. $C_0 = 1$ and $L_0 = 0.3$ nits = 0.9 lux on white). Formula (1) may also be interpreted in the sense that the product of G and D_0 is a general function of contrast and luminance.

In order to find out to what extent formula (1) applies to a given person, one may use this formula to calculate log G for each of the 40 threshold situations examined (4 luminosities × 10 contrast values). If the formula applies strictly, then the same value should be found in all 40 cases, or at least (owing to accidental measuring errors) a number of values grouped around an average. This appears to be the case with a great number of persons (the normal type). There are cases, however, in which the values found tend to increase or decrease according as C_0 or L_0 increases. Persons for which such results are found do not fit in completely with the general picture as represented by formula (1), but on the other hand the deviations are not so large as to invalidate the formula. In each special case the log G value to be substituted in formula (1) is taken as the average of the 40 individually calculated values quoted above. Finally, according to the behaviour of G, the test persons can be divided into a number of types as indicated in *table I*. Type A is normal and was found to comprise 67 % of all the persons tested.

Table I. Survey of visual types.

Type	Dependence of G on		Numbers of persons (%)
	increasing luminance	increasing contrast	
A	—	—	67
A	—	—	67
B_1	increase	—	11
B_2	decrease	—	5
C_1	—	increase	11
C_2	—	decrease	1
B_1C_1	increase	increase	4
B_2C_2	decrease	increase	1

The influence of age

It appears that the visual power G bears a simple relationship to the age of the test person. The test persons were divided into age groups and the average value of log G was calculated for each group. The mean value $\overline{G}$ thus obtained has been plotted in *fig. 6* against the average age $\overline{A}$ of the group. It can be derived from this graph that

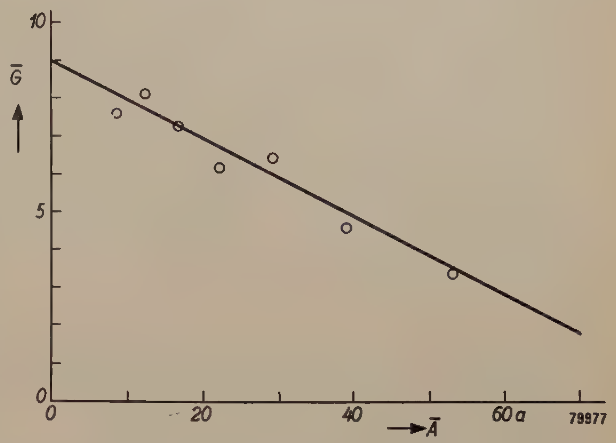


Fig. 6. The average value of the visual power $\overline{G}$ for various age groups, plotted against the average age $\overline{A}$ (in years) in these groups.

$$\overline{G} = 9 - 0.1\,\overline{A} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

in which A is expressed in years.

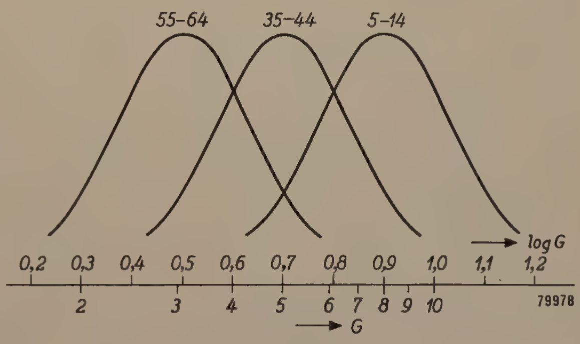


Fig. 7. Relative probability of G values for three different age groups. The figure was obtained as follows: For each age group it was calculated for how many persons the value of $\log G = g$ was situated in the intervals -0.025 to 0.025, 0.025 to 0.075, etc. From these data the average value of log G (g_m) for each age group (see fig. 6) was calculated, and also the standard deviation, i.e. the root mean square of the expression ($g - g_m$). These standard deviations amounted for six successive age groups to 0.095, 0.15, 0.11, 0.15, 0.28 and 0.14 respectively. For the sake of simplicity it has been assumed that in each group the log G values have a Gaussian distribution, and that for all groups the standard deviation $s_m = 0.16$,i.e. the mean of the values mentioned above. The curves drawn here can thus be expressed by the formula.

$$\exp\left[-(g - g_m)^2 / 2 s_m{}^2\right].$$

This formula has also been used to calculate the curves of figure 8.

---

[4] Here D_0 is expressed in minutes and the luminance L_0 in nits (cd/m²). In the articles mentioned in note [1]) the brightness (B_0) is expressed in millilamberts (1 mL = 3.18 nits).

[5] The author has suggested the introduction for G of a unit called "snellen", in honour of the Dutch opthalmologist Herman Snellen (1839-1918).

Naturally this empirical formula may be applied only within the age limits for which it has been derived, i.e. 10 to 60 years. It shows that visual acuity drops a by factor $^1/_2$ when going from the age of 10 to 50, and, after some slight extrapolation, by a further factor of $^1/_2$ from 50 to 70 years. That even considerable extrapolation does not lead to absurd results appears from the fact that according to formula (2) $\bar{G} = 0$ for $\bar{A} = 90$ years.

Of course there is a certain amount of dispersion for G in each age group. This dispersion is shown by *fig. 7*. *Figure 8* shows that in general the steep drop of G with increasing age, as represented in fig. 6 and by equation (2), is only slightly affected by this dispersion. Here the percentage of test persons in a given age group whose G value exceeds a given value is indicated. The figure proves that only the best 10% in the 55-64 years group ($\bar{A} = 60$) and the poorest 10% in the 5-14 years group ($\bar{A} = 10$) can see approximately as well as the average ones in the 35-44 year group ($\bar{A} = 40$).

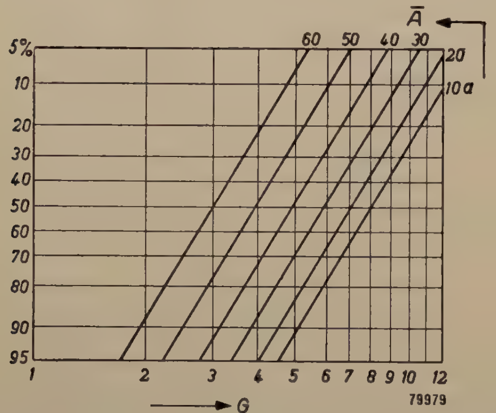


Fig. 8. Percentage of test persons for which the value of G exceeds the value plotted on the abscissa, with the average age $\bar{A}$ of the group as parameter.

The variations of D_0

When reviewing once more the results of the whole test, one is struck by the fact that the value of D_0, the smallest detail observed, can vary so strongly according to conditions. *Table II* gives an idea of this. It contains the D_0 values calculated with formula (1) for various contrast and luminance values and for various values of visual power G, viz. for the average value of G in the age groups 5 to 14 years, 25 to 34 years, and 45 to 54 years, and for the highest and lowest G values found ($G = 14$, and $G = 1$ respectively).

Apart from this selection of numerical values, some graphic representations of formula (1) are given in *figures 9* and *10*, in order to illustrate the influence of the various variable quantities. In fig. 9*a*, *b* and *c*, three age groups (the same as in table II) are dealt with separately. When applying the result to an individual person, account must be taken of the fact that, owing to the dispersion, the effective age [6]) of a person may differ from his real age.

In fig. 10, L_0 has been plotted against the effective age for three values of the contrast and for two values of D_0, viz. 1′ and 10′ (0.00027 and 0.0027 radians). It follows from this that, in order to discern the same detail with increasing age, the luminance must be more strongly increased for small details than for larger details. This explains why elderly people are inclined (erroneously) to forbid children to read in waning light, telling them

Table II. The smallest perceptible detail D_0 (in minutes of arc) for three contrast values C_0 and four luminance values L_0 at a visual power G = 1, 4, 6, 8 and 15.

	$G = 1$			$G = 4$			$G = 6$			$G = 8$			$G = 14$		
L_0(nit) \ C_0 =	0.1	0.32	1.0	0.32	1.0	1.0	0.32	1.0	0.1	0.32	1.0	0.1	0.32	0.32	1.0
1	26	13.5	2.7	6.5	3.4	4.1	2.1	1.15	1.15	3.3	1.7	0.9	1.8	0.96	0.5
10	13.5	9.1	4.9	3.4	2.0	12.5	1.2	1.3	0.8	1.7	1.0	0.6	0.95	0.6	0.35
100	9.5	5.8	3.7	22.2	1.4	0.95	1.4	0.9	0.6	1.1	0.7	0.45	0.6	0.4	0.25
1000	6.3	4.5	3.1	1.6	1.1	0.8	1.0	0.7	0.5	0.8	0.55	0.4	0.45	0.30	0.22

---

[6]) By effective age A_{eff} is understood the age a person should be in order to have the G value following from formula (2) if $\bar{G}$ were substituted by G, and $\bar{A}$ by A_{eff}, thus

$$A_{\mathrm{eff}} = 10(9 - G).$$

For G values over 9, A_{eff} is negative.

that they are "spoiling their eyes", because they are no longer able themselves to discern small type under such conditions.

If, on the other hand, the luminance is a given value, then fig. 9 gives the answer to the value of D_0 for the various age groups. The latter conforms with the situation met in television. In television the brightness of the picture can be adjusted to a certain extent by turning a knob, but there is an upper limit to this because, owing to the discontinuity of the image, flicker may occur, and because, even if the other conditions remain unchanged, the critical flicker frequency (i.e. the frequency above which no flicker is seen) is higher as the brightness is greater. With the fluorescent materials commonly used in Europe and at a mains frequency of 50 c/s, the maximum permissible screen luminance can be taken at 100 nits (100 cd/m^2) 7).

As the tests described above were carried out with the aid of Landolt rings and the details observed in a television image are of a different kind (brightness variations along the lines of the image) one cannot simply apply formula (1) to television. On the other hand the perception of details does not depend strongly on the nature of the detail. It is therefore still admissible to use formula (1) to get at least an impression of the variations in perception and their dependence on the various factors which affect them.

From Table II it can be seen that if $C_0 = 0.1$ and $L_0 = 100$ nits, D_0 may vary from 2.2′ ($G = 4$, $A_{\text{eff}} = 50$ years) to 1.1′ ($G = 8$, $A_{\text{eff}} = 10$ years), and in extreme cases even from 9.5 'to 0.6′. If we assume that formula (1) may be applied quantitatively to the case in question, this means that the distance *a* between the screen and the eye at which practically all details can be seen, varies from 1.5 metres ($A_{\text{eff}} = 10$ years) to 0.8 metre ($A_{\text{eff}} = 50$ years), and that in extreme cases *a* may even vary from 3.4 to 0.18 metres!

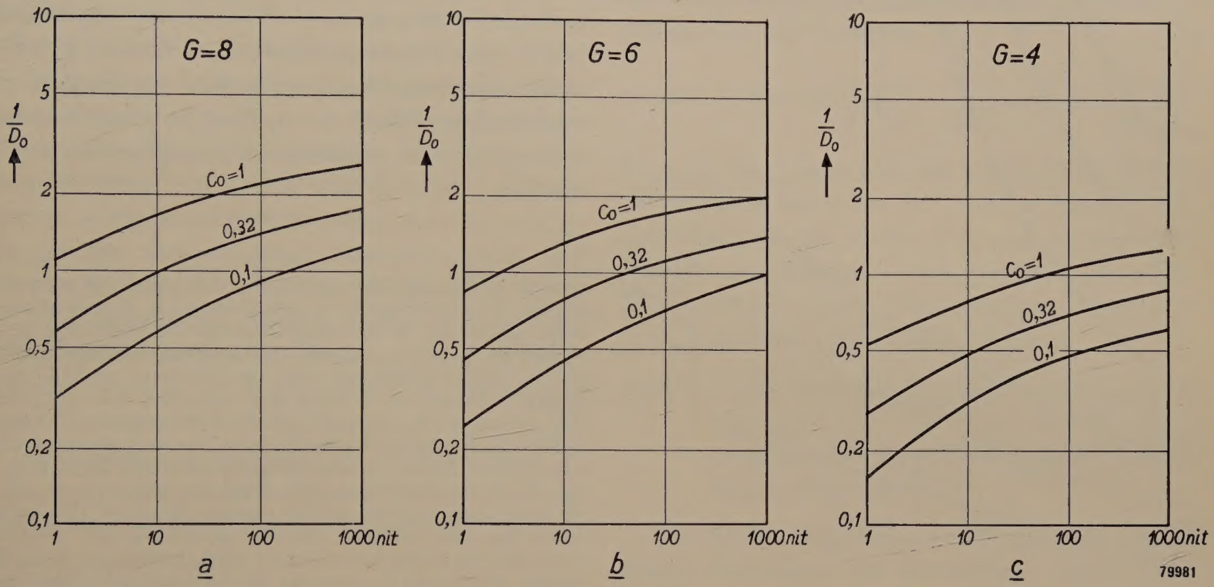


Fig. 9. The visual acuity ($1/D_0$) plotted on a double logarithmic scale against the luminance L_0 of the background, for three different values of the contrast ($C_0 = 0.1$, $C_0 = 0.32$, $C_0 = 1.0$). *a*) Age group 5-15 years; *b*) 25-35 years); *c*) 45-55 years.

The above conslusions can also be applied, mutatis mutandis, to the observation of a projected image (lantern slides, cinema projection). Here, too, with a view to the visibility of the details, it will be desirable to raise the luminance as much as possible, although here too there is an upper limit, this time set by the capacity of projector and lamp and the distance between projector and screen (in the case of cinema projection special measures are taken to avoid the inconvenience of flicker of the image 8)). It is a well-known fact that in general a screen lighting of 100 lux (screen luminance 20 to 100 nits, according to the reflection properties of the screen) is considered satisfactory. If we assume that the smallest details on a 4 × 3 m screen are 2.5 mm in size (this depends on the grain size of the negative with which the exposure was made and on the sharpness of the optical system of the camera) we

---

7) See J. Haantjes and F. W. de Vrijer, Flicker in television pictures, Philips tech. Rev. **13**, 55-60, 1951/52.

8) This is explained on p. 161 of the article by J. Kotte, A professional cine projector for 16 mm film, in this issue.

find that the optimum distance between observer and screen varies from 4 to 8 metres according to the observer's age, and from 1 to 18 metres if allowance is made for extreme cases.

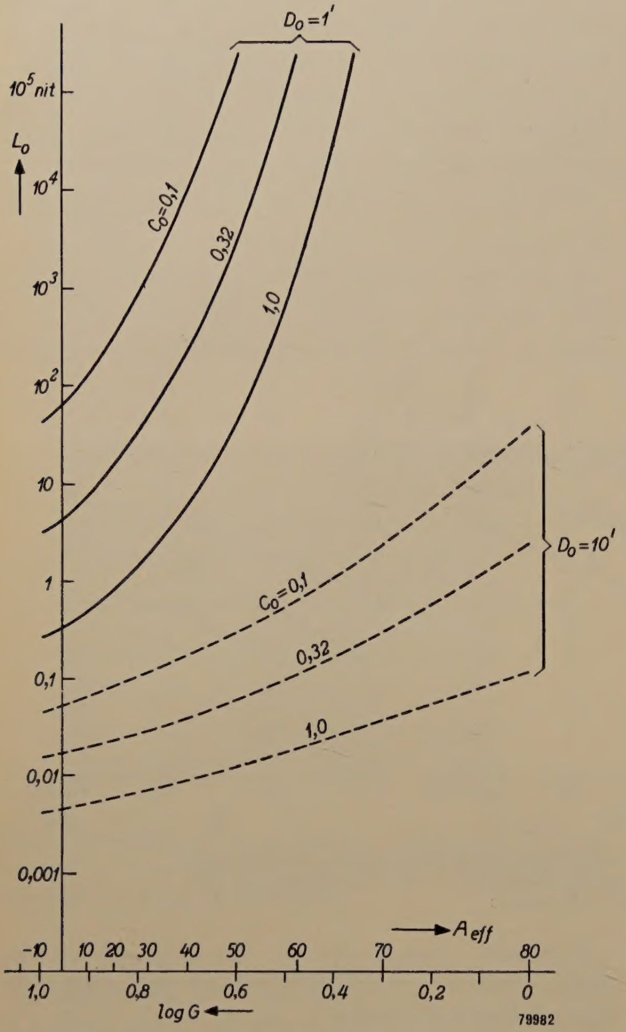


Fig. 10. Luminance L_0 required for perception of a detail $D_0 = 1'$ and $10'$ respectively at contrasts C_0 of 0.1, 0.32 and 1.0, as a function of the effective age A_{eff}.

We may conclude by pointing out that formula (1) should be used with some care for calculating lighting standards. In the first place the tests refer to threshold values, indicating a minimum which in practice should always be exceeded. It is difficult to say to what extent the minimum should be surpassed, because apart from perception the comfort of the subject plays also a part. Furthermore, the formula is the outcome of laboratory experiments, where *time* is not included as a factor. It will be clear, therefore, that conclusions drawn from such experiments should be applied to cases involving movement and rapidity of perception with great caution.

Summary. In the observation of a television (or projection) image, the visual acuity of the observer is of importance. A measure for this is the smallest detail D_0 which can be discerned at a given contrast and a given luminance of the object. Tests carried out with Landolt rings on 228 persons in ages varying from 7 to 64 years revealed that the product GD_0 is a general function of contrast and luminosity, G being an individual constant (the visual power) for each observer, whose value, apart from a certain statistical variation depends on the age A of the observer in the following way: $G = 9 - 0.1\,A$. A description is given of the manner in which this average result has been obtained, and the restrictions are mentioned that must be taken into account, when applying it to an individual observer. The meaning of the relationship thus found is illustrated by a table and several graphs. It appears that the visual acuity increases by a factor of roughly 2, if a) the luminance increases by a factor 100, or b) the contrast increases by a factor 5. Under equal conditions the visual acuity of persons 50 years old is half that of a ten-year old child, while that of persons 70 years old is only one quarter. Although in the case of details of different kinds it is not permissible to draw quantitative conclusions from the relationship found, it still helps to give an impression of the extent to which the need for light increases with age and of the extent to which, at a given luminance and contrast, the perception of details is better with young persons than with old.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the administration of the Philips Research Laboratory, Eindhoven, Netherlands.

2099: H. J. G. Meyer and D. Polder: Note on polar scattering of conduction electrons in regular crystals (Physica **19**, 255-264, 1953).

The reciprocal relaxation time of conduction electrons due to polar interaction with piezoelectrically active acoustical vibration modes is calculated for two-atomic regular crystals. It is found to be proportional to $T^{1/2}$. It is shown for the case of ZnS (sphalerite) that below 150 °K the numerical value of this reciprocal relaxation time is comparable with the reciprocal relaxation time due to polar interaction with optical vibration modes. In crystals with sodium chloride structure the mass difference of the ions may in principle be responsible for polar scattering by acoustical modes. The corresponding reciprocal relaxation time is proportional to $T^{3/2}$. In the special case of LiF it becomes equal to polar scattering by optical modes at about 200 °K.

2100: H. A. Klasens: On the nature of fluorescent centres and traps in zinc. sulphide (J. Electrochem. Soc. **100**, 72-80, 1953, No. 2).

The impurity levels in the energy diagram of a zinc sulphide phosphor are considered to be localized S^{2-} levels lifted above the filled S^{2-} band due to the presence of monovalent positive or trivalent negative activator ions in the lattice. Electron traps are formed similarly by the substitution of S^{2-} ions by monovalent negative ions or of Zn^{2+} ions by trivalent positive ions. The energy produced when electrons recombine with trapped holes or when holes recombine with trapped electrons is either emitted directly as light or is first transferred to impurity ions. The elements of the iron group give rise to electron traps. The killing action of these elements is explained by assuming that the energy liberated by recombination between holes and electrons in these traps is transferred to the killer ions. The excited ions return to the ground state without radiation, because of the presence of many electronic levels between the excited and the ground state. The effect of heat and infrared radiation on the luminescence is discussed. It is shown that, in a phosphor, energy may be transferred by electrons through the conduction band or by holes through the occupied S^{2-} band.

2101: J. I. de Jong: The determination of formaldehyde in the presence of phenols and phenol alcohols (Rec. Trav. chim. Pays-Bas **72**, 356-357, 1953).

A method described by Pfeil and Schroth for the quantitative determination of formaldehyde appears to be applicable in the presence of phenols and phenol alcohols.

2102*: P. Cornelius: L'électricité selon le système Giorgi rationalisé (Ed. Dunod, Paris 1953, 116 pp.). (Electromagnetism according to the rationalised Giorgi system: in French.)

French translation of Dutch text (see these abstracts, No. **1803**). The author has extended and altered the text at a few points. The generator (dynamo) is treated from the viewpoint of an observer moving with the rotor (alternating flux) as well as from that of an observer at rest (Lorentz force on electrons in wire). The case of a beam of electrons in a magnetic field is considered. In the final chapter the author explains why he avoided discussions on quantity equations. (The equations in Giorgi units as given in the text may be considered either as numerical or as quantity equations.)

2103: H. Bremmer Eine einfache Näherungsformel für die Feldverteilung längs der Achse magnetischer Elektronenlinsen mit ungesättigten Polschuhen (Optik **10**, 1-4, 1953, No. 1). (A simple approximate formula for the field along the axis of magnetic electron lenses with unsaturated pole-pieces; in German.)

For the field $H(z)$ along the axis of a rotationally symmetrical magnetic electron lens with unsaturated pole-pieces, the formula

$$H = (0.4\,\pi)\,NI\,\left\{\varphi\left(\frac{z}{b}+\frac{s}{2b}\right)-\varphi\left(\frac{z}{b}-\frac{s}{2b}\right)\right\}\Big/ s$$

is given, in which s is the width of the slit, b the diameter of the bore, and $\varphi(x)$ a function which, depending on the value of x, can be expanded in a series in different ways. A table of $\varphi(x)$ for $0 < x < \infty$ is given. The results are in agreement with a (less simple) theory of Lenz.

2104*: J. A. Haringx: Stresses in corrugated diaphragms (pp. 198-213, C. B. Bienezo Anniversary Volume on Applied Mechanics, Stam, Haarlem 1953).

The stress components in the material of a corrugated diaphragm are determined, approximately, by replacing the diaphragm by a ficticious flat plate of similar properties as described in a previous paper. From the most unfavourable combination of these stress components that is ever possible, formulae for the maximum stresses were derived. These formulae were tested in a special case for which the stresses have been worked out in full by Grover and Bell, and are found to be in good agreement. Thus these formulae may be useful for design purposes.

R 216: F. van der Maesen, P. Penning and A. van Wieringen: On the thermal conversion of germanium (Philips Res. Rep. **8**, 241-244, 1953, No. 4).

Experiments were carried out regarding the conversion of *n*-type germanium into *p*-type by heat treatment at 800 °C. The results indicate that the conversion is due to the presence of acceptor impurities on the Ge surface prior to heating. As long as the surface is clean, its roughness seems to have no influence on the conversion. Experiments with pieces of Ge saturated by diffusion with Cu at 800 °C show that the Cu can be inactivated by heat treatment at 500 °C but that it remains present in some form throughout the material. A similar effect was found with Ni.

R 217: W. K. Westmijze: Studies on magnetic recording, III (Philips Res. Rep. **8**, 245-269, 1953, No. 4).

Continuation of **R 213** end **R 214**. The recording methods leading to a linear relationship between the input signal and the recorded magnetization are discussed. Special attention is paid to the a.c. biasing method and its relation to ideal magnetization. This discussion is based on magnetic measurements in homogeneous fields. A magnetic model explaining the linearizing effect of the a.c. biasing field is discussed.

Next follows the calculation of the magnetic field that exists in and around a sinusoidally magnetized tape, the cases of longitudinal and perpendicular magnetization being treated separately. When the permeability of the tape is greater than unity, the demagnetizing field in the tape effects a decrease of the recorded magnetization. The flux in an ideal reproducing head is calculated for this case. It is shown that longitudinal and perpendicular magnetization produce the same flux in the head only when the permeability of the tape is equal to 1.

R 218: J. L. Meijering: Interface area, edge length, and number of vertices in crystal aggregates with random nucleation (Philips Res. Rep. **8**, 270-290, 1953, No. 4).

The interface area, edge length, and numbers of faces, edges and vertices in an aggregate consisting of a large number of crystals are calculated for two models. In the first ("cell model") the crystals start to grow simultaneously and isotropically from nuclei distributed at random. In the second ("Johnson-Mehl" model) the nuclei appear at different moments, the rate of nucleation being constant. Corresponding calculations are made for plane sections of the aggregates and for two-dimensional aggregates. For the one-dimensional case the size-distribution curves are calculated. From a discussion of the results it is concluded that in the two- and three-dimensional cell models, the crystals are less equiaxial than in the corresponding Johnson-Mehl models.

R 219: A. H. Boerdijk: The value of the constant in Wien's displacement law (Philips Res. Rep. **8**, 291-303, 1953, No. 4).

It is shown that the current definition of monochromatic intensity (I_λ) contains an arbitrary element. Other definitions are considered, based on a logarithmic wavelength scale (I_r) or on a frequency scale (I_ν). For black-body radiation the maxima of these intensities occur at different places in the spectrum, giving rise to three different constants in Wien's law. The question arises which scale should be used. It appears that the use of I_λ is a matter of convention only. There are logical arguments for using I_r, the corresponding Wien-constant being 0.3668 cm °K.

R 220 and **R 221:** F. C. Romeijn: Physical and crystallographical properties of some spinels (Philips Res. Rep. **8**, 304-342, 1953, Nos. 4 and 5).

From X-ray measurements on simple and complicated spinels, regularities in the ionic distribution and lattice constants have been investigated. These regularities have been explained partly by general methods (Madelung potential, geometrical considerations) and partly by also taking into account the individual properties of the ions that are correlated with the distribution of electrons within the ion. The calculated correlation between the ionic distribution and the oxygen parameter u was found to be confirmed by experiment. The ultimate choice of the distribution, however, is governed by the individual properties of the ions, partly by their dimensions and partly by the distribution of electrons. Some physical properties of the compounds investigated have been correlated with the ionic distribution.